

Optimization of the Variational Quantum Eigensolver for Quantum Chemistry Applications

Citation for published version (APA):

de Keijzer, R. J. P. T., Colussi, V. E., Škorić, B., & Kokkelmans, S. J. J. M. F. (2021). Optimization of the Variational Quantum Eigensolver for Quantum Chemistry Applications. *arXiv*, 2021, [2102.01781]. <https://arxiv.org/abs/2102.01781>

Document status and date:

Published: 02/02/2021

Document Version:

Publisher's PDF, also known as Version of Record (includes final page, issue and volume numbers)

Please check the document version of this publication:

- A submitted manuscript is the version of the article upon submission and before peer-review. There can be important differences between the submitted version and the official published version of record. People interested in the research are advised to contact the author for the final version of the publication, or visit the DOI to the publisher's website.
- The final author version and the galley proof are versions of the publication after peer review.
- The final published version features the final layout of the paper including the volume, issue and page numbers.

[Link to publication](#)

General rights

Copyright and moral rights for the publications made accessible in the public portal are retained by the authors and/or other copyright owners and it is a condition of accessing publications that users recognise and abide by the legal requirements associated with these rights.

- Users may download and print one copy of any publication from the public portal for the purpose of private study or research.
- You may not further distribute the material or use it for any profit-making activity or commercial gain
- You may freely distribute the URL identifying the publication in the public portal.

If the publication is distributed under the terms of Article 25fa of the Dutch Copyright Act, indicated by the "Taverne" license above, please follow below link for the End User Agreement:

www.tue.nl/taverne

Take down policy

If you believe that this document breaches copyright please contact us at:

openaccess@tue.nl

providing details and we will investigate your claim.

Optimization of the Variational Quantum Eigensolver for Quantum Chemistry Applications

R.J.P.T. de Keijzer,¹ V.E. Colussi,^{1,2} B. Škorić,¹ and S.J.J.M.F. Kokkelmans¹

¹*Eindhoven University of Technology, P. O. Box 513, 5600 MB Eindhoven, The Netherlands**

²*INO-CNR BEC Center and Dipartimento di Fisica, Università di Trento, 38123 Povo, Italy*

(Dated: February 4, 2021)

This work studies the variational quantum eigensolver algorithm, designed to determine the ground state of a quantum mechanical system by combining classical and quantum hardware. Methods of reducing the number of required qubit manipulations, prone to induce errors, for the variational quantum eigensolver are studied. We formally justify the qubit removal process as sketched by Bravyi, Gambetta, Mezzacapo and Temme [arXiv:1701.08213 (2017)]. Furthermore, different classical optimization and entangling methods, both gate based and native, are surveyed by computing ground state energies of H_2 and LiH. This paper aims to provide performance-based recommendations for entangling methods and classical optimization methods. Analyzing the VQE problem is complex, where the optimization algorithm, the method of entangling, and the dimensionality of the search space all interact. In specific cases however, concrete results can be shown, and an entangling method or optimization algorithm can be recommended over others. In particular we find that for high dimensionality (many qubits and/or entanglement depth) certain classical optimization algorithms outperform others in terms of energy error.

I. INTRODUCTION

Presently, quantum computing is in the noisy intermediate-scale (NISQ) era [1], where the available universal quantum computers cannot outperform their classical counterparts except for a few, specific, well-designed cases [2]. However, even in the NISQ era quantum computers can be used in practice. One such application is the variational quantum eigensolver (VQE) algorithm, first proposed in a paper by Peruzzo and McClean in 2014 [3]. The algorithm is hybrid, e.g. at certain points a quantum processor unit (QPU) is addressed. The VQE algorithm has shown proof of concept for small molecules with several different designs of qubits [4–8]. The aim of the algorithm is to determine the lowest eigenvalue of a quantum Hamiltonian, which can be equated to finding the ground state energy of a molecule. The algorithm has also been implemented to solve NP-hard problems, such as the travelling salesman problem [9], in polynomial time.

To illustrate how the VQE algorithm works, a schematic of the relevant circuit diagram is shown in Fig. 1. To determine the ground state energy of a Hamiltonian the VQE algorithm maps the fermionic degrees of freedom of molecules onto a set of qubits. These qubits can then be manipulated into a trial state with a certain energy by applying a sequence of entanglement operators U_{ent} and individual rotations $U_{i,j}$. The length of this sequence is the depth d which plays an important role in determining which part of the state space can be addressed by the VQE. The entanglement operator U_{ent} is a multi-qubit gate that ensures entangled states

can be reached and thus its choice is critical in exploiting the quantum nature of the method. This operator is necessary for utilizing the computational powers of the multi-qubit structure. By applying a problem dependent Hamiltonian to the trial state, its energy can be measured using a QPU. Based on these measurements the classical part of the computer proposes a new trial state (new rotations) by means of a classical optimization algorithm [10]. Because of the limited hardware of quantum computers, each qubit manipulation has a non-negligible probability of error. Therefore, it is important to minimize the number of necessary manipulations.

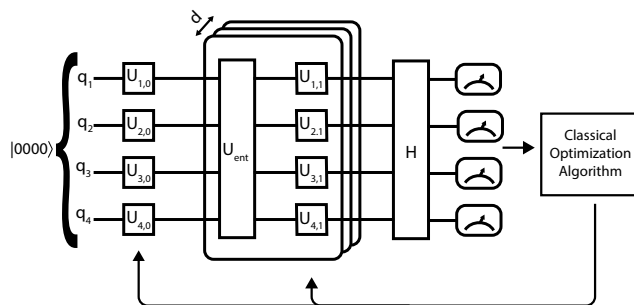


Figure 1: Schematic circuit diagram of the VQE algorithm. The qubits are manipulated by a sequence of entanglement operators U_{ent} and individual rotations $U_{i,j}$ into a prepared trial state.

This paper examines three ways of achieving this minimization: (i) by reducing the number of qubits used to describe the problem as proposed by Bravyi et al. [11], (ii) by reducing the number of qubit manipulations using entanglement methods native to the physical system, and (iii) by optimizing the classical search algorithm used to converge on the ground state.

* Corresponding author: r.j.p.t.d.keijzer@student.tue.nl

The layout of this paper is as follows. In Secs. II and III the VQE algorithm and the method of simulation are described. Sec. IV gives a constructive proof of the scheme by Bravyi et al. [11] Secs. V and VI respectively describe multi-qubit entangling methods and classical optimization algorithms. Finally, in Sec. VII we compare and contrast the methods and their results on small molecule VQE problems.

II. QUANTUM MEASUREMENTS AND SIMULATION SETUP

The VQE algorithm exploits the variational principle and aims to find a set of parameters minimizing the energy of a quantum system. The states of a system are mapped to qubits and the classical optimization algorithm optimizes the variational calculation. To do so for Hamiltonians of small molecules requires a characterization of their electron spin-orbitals. For this the Hartree-Fock method is applied, which reformulates the Hamiltonian in first-quantization form [12]. The Born-Oppenheimer approximation is applied which ensures that the nuclear contribution to the Hamiltonian is constant [12]. The Hamiltonian H_1 under the Born-Oppenheimer approximation becomes

$$H_1 = - \sum_i \frac{\nabla_{\vec{r}_i}^2}{2} - \sum_{i,j} \frac{Z_i}{|\vec{R}_i - \vec{r}_j|} + \sum_{i,j>i} \frac{1}{|\vec{r}_i - \vec{r}_j|} + \sum_{i,j>i} \frac{Z_i Z_j}{|\vec{R}_i - \vec{R}_j|}, \quad (1)$$

where, \vec{R}_i and Z_i are the position and charge of nucleus i . \vec{r}_i is the position of electron i . In VQE, the Hamiltonian is described in Hartree units [13]. The Hamiltonian is reformulated in second-quantization form [14] H_2 , by projecting it onto a finite set of orthogonal spin-orbital modes $\{\phi_1\}_{i=1}^n$, as

$$H_2 = V_{nn} + \sum_{p,q} h_{pq} a_p^\dagger a_q + \frac{1}{2} \sum_{p,q,r,s} h_{pqrs} a_p^\dagger a_q^\dagger a_r a_s, \quad (2)$$

where the nuclei-nuclei interaction potential $V_{nn} = \sum_{i,j>i} Z_i Z_j / |\vec{R}_i - \vec{R}_j|$ is constant and the annihilation and creation operators a_i^\dagger and a_i work on a set of orthogonal spin-orbital modes $\{\phi_1\}_{i=1}^n$ determined by the linear combination of atomic orbitals (LCAO) method and which induce the fermionic algebra [15]. The number of considered orbitals determines the complexity of the problem and with that the number of required qubits. For this paper the $1s$ orbital for hydrogen H_2 and the $1s$, $2s$ and $2p_x$ orbitals for lithium-hydrogen LiH (aligned on the x -axis) are low enough in energy to be considered in the LCAO construction. All higher energy atomic orbitals are ignored. The molecular integrals h_{pq} and h_{pqrs}

are referred to as the one-electron and two-electron integrals respectively

$$h_{pq} = \int S_p^*(\omega) S_q(\omega) d\omega \times \int_{\mathbb{R}^3} \phi_p^*(\vec{r}) \left(\frac{\nabla_{\vec{r}}^2}{2} - \sum_i \frac{Z_i}{|\vec{R}_i - \vec{r}|} \right) \phi_q(\vec{r}) d\vec{r}, \quad (3)$$

$$h_{pqrs} = \int \int S_p^*(\omega_1) S_s(\omega_1) S_q^*(\omega_2) S_r(\omega_2) d\omega_1 d\omega_2 \times \int_{\mathbb{R}^3} \int_{\mathbb{R}^3} \frac{\phi_p^*(\vec{r}_1) \phi_q^*(\vec{r}_2) \phi_r(\vec{r}_2) \phi_s(\vec{r}_1)}{|\vec{r}_1 - \vec{r}_2|} d\vec{r}_1 d\vec{r}_2, \quad (4)$$

where $S_i(\omega)$ is the spin part of the spin orbitals [16].

Some important concepts in quantum computing are introduced below. These are required to understand the VQE method and are used in proofs that follow.

Definition 1. Pauli Matrices

$$I := \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}, \quad X := \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \quad Y := \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix}, \quad Z := \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}. \quad (5)$$

Definition 2. m -fold Single Qubit Operator

$$\sigma_j^\rho := I^{\otimes(j-1)} \otimes \rho \otimes I^{\otimes(m-j)}, \quad j \in 1, \dots, m, \quad \rho \in \{X, Y, Z\}. \quad (6)$$

These m -fold single qubit operator work on a m -qubit state but only change the state of one of the qubits.

Definition 3. m -fold Pauli operator space P_m

$$P_m := \{\pm 1, \pm i\} \times \{I, X, Y, Z\}^{\otimes m}. \quad (7)$$

P_m forms a group under operator multiplication since the Pauli operators P_1 form a group. Note that since the Pauli matrices are a complete set for the Hilbert space of complex 2×2 matrices, P_m is a complete set for the Hilbert space of complex $2^m \times 2^m$ matrices

The Jordan-Wigner transformation can be used to map the creation and annihilation operators of Eq. (2) to combinations of Pauli matrices [17]. The transformation is given by

$$a_j = I^{\otimes j-1} \otimes \sigma^+ \otimes Z^{\otimes m-j}, \quad (8)$$

$$a_j^\dagger = I^{\otimes j-1} \otimes \sigma^- \otimes Z^{\otimes m-j}, \quad (9)$$

where m is the number of qubits. The σ^\pm operators are defined as

$$\sigma^\pm := \frac{X \pm iY}{2}. \quad (10)$$

It can be shown that the transformed operators a_j and a_j^\dagger obey the canonical commutation relations for

fermions. Since this transformation is local the electron configuration can be determined from the qubit state. Since the number of required qubits is a key factor determining the required effort for calculation on both physical systems and in simulations it is desirable to reduce the number of qubits where possible. This can be done by projecting only on those states describing the right number of electrons (e.g. for LiH the state $|101100\rangle$ would describe three electrons while only two electrons, in the outer shells, are analyzed for the problem). For this purpose a projector operator Π_{elecN}^m on the states describing N electrons is defined as

$$\Pi_{elecN}^m := \prod_{j=0, \dots, m}^{j \neq N} \frac{N_{state} - j}{N - j}, \quad (11)$$

where N_{state} is the operator returning the state it operates on multiplied by the number of electrons described by the state. A similar projection can be done on spin up and down of states. Another way of reducing the number of qubits necessary to describe the problem is using the symmetry of the Hamiltonian as described in Sec. IV.

Because P_m is a complete set for the Hilbert space of complex $2^m \times 2^m$ matrices, any m -qubit Hamiltonian H can be written as a linear combination of elements of the m -fold Pauli operators in P_m as

$$H = \sum_k h_k \sigma^k, \quad \sigma^k = \sigma_1^k \otimes \sigma_2^k \otimes \dots \otimes \sigma_m^k \in P_m,$$

$$h_k = \langle H, \sigma^k \rangle_{\mathcal{F}} = \sum_{i,j} \overline{H_{ij}^T} \sigma_{ij}^k = \text{tr}(H^\dagger \sigma^k). \quad (12)$$

This is a finite dimensional Fourier decomposition with the Frobenius inner product $\langle \cdot, \cdot \rangle_{\mathcal{F}}$ on matrix representations of the operators. Note that in Eq. (12) σ_{ij}^k is a matrix element, not a Pauli matrix working on a qubit. Any rotation on a single qubit can be written as

$$R(\alpha, \beta, \gamma, \delta) = e^{i\alpha} Z_\beta X_\gamma Z_\delta, \quad (13)$$

where $\alpha, \beta, \gamma, \delta \in [0, 2\pi]$. The operators X_ϕ and Z_ϕ denote a rotation of an angle ϕ around the given axis and are given by

$$\begin{aligned} X_\phi &= \cos(\phi/2)I + \sin(\phi/2)X, \\ Y_\phi &= \cos(\phi/2)I - i \sin(\phi/2)Y, \\ Z_\phi &= \cos(\phi/2)I + \sin(\phi/2)Z. \end{aligned} \quad (14)$$

III. TRIAL STATE INITIALIZATION AND MEASUREMENT

In order to initialize a trial state the individual qubits have to be put in a requested state using a sequence of rotation and entanglement operations. The depth d of a state preparation is defined as the length of this sequence, or equivalently the number of entanglement operations applied. Each state on the Bloch sphere of a single qubit can be reached with a ZXZ -rotation. Such a rotation on a qubit q at a depth i can be written as

$$U_{q,i}(\vec{\theta}) = Z_{\theta_1^q, i} X_{\theta_2^q, i} Z_{\theta_3^q, i}, \quad (15)$$

where $\vec{\theta}$ has three elements for every qubit and depth pair. In total $3d + 3$ rotations would be done for every qubit. However, in this paper the ansatz state will always be the vacuum state. The first Z -rotation can therefore be omitted. Thus the parameter vector $\vec{\theta}$ is an element of the search space $[0, 2\pi]^{\otimes D}$, where $D = (3d + 2)m$. To reach the trial state described by the parameter vector $\vec{\theta}$ the ansatz state $|\psi_{init}\rangle$ is rotated and entangled. The entanglement operator is U_{ent} . The prepared trial state after all rotations and entanglements will be

$$\begin{aligned} |\Psi(\vec{\theta})\rangle &= \left(\prod_{q=1}^m U_{q,d}(\vec{\theta}) \times U_{ent} \right) \times \left(\prod_{q=1}^m U_{q,d-1}(\vec{\theta}) \times U_{ent} \right) \times \dots \\ &\times \left(\prod_{q=1}^m U_{q,0}(\vec{\theta}) \right) |\psi_{init}\rangle. \end{aligned} \quad (16)$$

With the trial state prepared, the expectation value of H can be measured on a quantum computer as

$$\begin{aligned} \langle H \rangle_{\vec{\theta}} &= \langle \Psi(\vec{\theta}) | H | \Psi(\vec{\theta}) \rangle = \sum_k h_k \langle \Psi(\vec{\theta}) | \sigma^k | \Psi(\vec{\theta}) \rangle \\ &= \sum_k h_k \langle \Psi(\vec{\theta}) | \sigma_1^k \otimes \sigma_2^k \otimes \dots \otimes \sigma_m^k | \Psi(\vec{\theta}) \rangle. \end{aligned} \quad (17)$$

In quantum computers these expectation values require sufficient measurements. In simulation the expectation values require an inner product of matrices and vectors.

IV. \mathbb{Z}_2 SYMMETRY QUBIT REDUCTION SCHEME

This section formally justifies the qubit removal process as described by Bravyi et al. [11]. In this process the \mathbb{Z}_2 symmetries in Hamiltonians [11], which often originate from geometric symmetries in the molecule, are exploited to reduce the number of qubits necessary to describe the Hamiltonian. The reductions will substantially lower computation power and time. The main idea of the qubit reduction scheme is to transform a m -qubit Hamiltonian in such a way that the Pauli operators in the decomposition will all have either I or X as the last $r \in \mathbb{N}$ factors (e.g. YXXZIX and ZXYZXX). If the i -th term of every Pauli operator in the decomposition is I or X then $[H, \sigma_i^x] = 0$, so the Hamiltonian H and σ_i^x commute and thus share common eigenvectors. One can therefore replace the i -th factors of the operators with the eigenvalues ± 1 and thus taper off a qubit [11]. We ignore the phases of the m -fold Pauli operators because these do not influence commutation. Doing so every m -fold Pauli operator becomes its own inverse. First some definitions are given and lemmas are proven in order to help provide a constructive proof of the scheme.

Definition 4. Symmetry Group

A symmetry group S of a group G is an abelian subgroup of G such that $-I \notin S$.

Definition 5. Center and Centralizer

Let G be a group. The center $Z(G)$ of G is the set of elements commuting with every element of G .

$$Z(G) = \{z \in G \mid zg = gz \quad \forall g \in G\}. \quad (18)$$

The centralizer $C_G(A)$ of $A \subseteq G$ is the set of all elements in G which commute with every element of A .

$$C_G(A) = \{g \in G \mid ag = ga \quad \forall a \in A\}. \quad (19)$$

Lemma 1. Symmetry group commuting with $2^m \times 2^m$ Hamiltonian

Let $S \subseteq P_m$ be a symmetry group. S commutes with a $2^m \times 2^m$ Hamiltonian $H = \sum_k h_k \sigma^k$ if S commutes with every Pauli operator in the decomposition of H . Thus, let $H = \sum_k h_k \sigma^k$. If $\forall k$ S commutes with σ^k then S commutes with H .

Proof. $sH = s \sum_k h_k \sigma^k = \sum_k h_k s \sigma^k = \sum_k h_k \sigma^k s = Hs$ \square

Lemma 2. Redefining of generators

Let the generators $\tau_1, \tau_2, \dots, \tau_r$ generate the abelian group $S = \langle \tau_1, \tau_2, \dots, \tau_r \rangle \subset P_m$ and let $i, j \in \{1, \dots, r\}$ with $i \neq j$. Then $S = \langle \tau_1, \tau_2, \dots, \tau_i, \dots, \tau_j, \dots, \tau_r \rangle = \langle \tau_1, \tau_2, \dots, \tau_i, \dots, \tau_{j-1}, \tau_i \tau_j, \tau_{j+1}, \dots, \tau_r \rangle$

Proof. Since $\tau_i, \tau_j \in S$ it holds that $\tau_i \tau_j \in S$. Therefore, $\langle \tau_1, \tau_2, \dots, \tau_i, \dots, \tau_{j-1}, \tau_i \tau_j, \tau_{j+1}, \dots, \tau_r \rangle \subseteq \langle \tau_1, \tau_2, \dots, \tau_i, \dots, \tau_j, \dots, \tau_r \rangle$. Now if $\tau_i \tau_j, \tau_i \in S$ then $\tau_i \tau_i \tau_j = \tau_j \in S$. Therefore $\langle \tau_1, \tau_2, \dots, \tau_i, \dots, \tau_j, \dots, \tau_r \rangle \subseteq \langle \tau_1, \tau_2, \dots, \tau_i, \dots, \tau_{j-1}, \tau_i \tau_j, \tau_{j+1}, \dots, \tau_r \rangle$. \square

Lemma 3. Commutation of product

Let A, B_1, B_2 and B_3 be operators such that A commutes with B_1 and anti-commutes with B_2 and B_3 . Then $B_1 B_2$ and $B_2 B_1$ anti-commute with A . $B_2 B_3$ and $B_3 B_2$ commute with A .

Proof. $AB_1 B_2 = B_1 A B_2 = -B_1 B_2 A$ and $AB_2 B_1 = -B_2 A B_1 = -B_2 B_1 A$. So indeed $B_1 B_2$ and $B_2 B_1$ anti-commute with A . Furthermore, $AB_2 B_3 = -B_2 A B_3 = B_2 B_3 A$ and $AB_3 B_2 = -B_3 A B_2 = B_3 B_2 A$. So indeed $B_2 B_3$ and $B_3 B_2$ anti-commute with A . \square

An important result of stabilizer theory [18] is that a symmetry group commuting with a m -qubit Hamiltonian will at most require m generators. This is a key fact in the reduction scheme (the exact proof of this is given by Fujii [19]). A proof sketch would take the following form. Each symmetry group commuting with a m -qubit Hamiltonian is a symmetry group of P_m . Each symmetry group generator projects a m -qubit basis state onto one of its eigenvalues ± 1 . Therefore, each of the generators divides the Hilbert space of m -qubits into two subspaces. The generators share a common set of eigenvectors as they all commute. The generators thus partition the m -qubit space into $2^{\#\text{generators}}$ subspaces. Since there are only 2^m basis states, $\#\text{generators} \leq m$.

Now the algorithm for the qubit reduction is given below. For notation purposes the single qubit operators X, Y and Z on qubit i are sometimes denoted as σ_i^X, σ_i^Y and σ_i^Z respectively. The first goal is to find a symmetry group S , commuting with H , which has a maximal number of generators. First off, every Pauli operator $\sigma^k \in P_m$ will be encoded as a binary row-vector $(a_x | a_z)$ of length $2m$ as follows

$$\sigma(a_x | a_z) = \sigma(a_{x1}, \dots, a_{xm} | a_{z1}, \dots, a_{zm}) = \prod_{i=1}^m (\sigma_i^x)^{a_{xi}} (\sigma_i^z)^{a_{zi}}, \quad (20)$$

where $a_{xi}, a_{zi} \in \{0, 1\}$. As an example $IZXY = (0011|0101)$. Using the standard inner product on $2m$ -dimensional vectors in \mathbb{Z}_2 , resulting in a value 0 or 1, the following commutation relation holds

$$\sigma(a_x | a_z) \sigma(b_x | b_z) = (-1)^{a_x \cdot b_z + a_z \cdot b_x} \sigma(b_x | b_z) \sigma(a_x | a_z). \quad (21)$$

Following this, a scalar product \times is defined on the vector space of binary vectors of length $2m$ as

$$a \times b := a_x \cdot b_z + a_z \cdot b_x \in \{0, 1\}. \quad (22)$$

This vector product is symmetrical and linear in both a and b . By evaluating this vector product it can easily be determined if the corresponding operators will commute or anti-commute, see Eq. (21).

The set of k Pauli operators $\sigma^1, \dots, \sigma^k$ appearing in the decomposition of H can be represented by vectors

$(a_x^1|a_z^1), \dots, (a_x^k|a_z^k)$ which can in turn be represented in a $k \times 2m$ matrix G as

$$G = \left[\begin{array}{ccc|ccc} \cdots & a_x^1 & \cdots & \cdots & a_z^1 & \cdots \\ \cdots & a_x^2 & \cdots & \cdots & a_z^2 & \cdots \\ & \vdots & & & \vdots & \\ \cdots & a_x^r & \cdots & \cdots & a_z^r & \cdots \end{array} \right] \quad (23)$$

Now if a Pauli operator $\sigma(b_x|b_z)$ were to commute with every term in H it must be that $G \times (b_x|b_z) = 0$, where the vector product \times , defined in Eq. (22), is taken for every row in G . From the theory of stabilizer codes the concept of a parity check matrix E is borrowed [20]. This matrix E is defined as

$$E = \left[\begin{array}{ccc|ccc} \cdots & a_x^1 & \cdots & \cdots & a_x^1 & \cdots \\ \cdots & a_x^2 & \cdots & \cdots & a_x^2 & \cdots \\ & \vdots & & & \vdots & \\ \cdots & a_x^r & \cdots & \cdots & a_x^r & \cdots \end{array} \right]$$

Note that $EG^T = 0$. Now if the operator $\sigma(b_x|b_z)$ were to be part of the symmetry group S then $(b_x|b_z) \in \ker(E)$. Next a basis for $\ker(E)$ can be constructed. Let $d := \dim(\ker(E))$. These basis vectors b^1, \dots, b^d give rise to Pauli operators $\sigma^1, \dots, \sigma^d \in P_m$. However, it is not guaranteed that these Pauli operators mutually commute and thus will form a set of generators for a symmetry group S . Commutation can be enforced by applying the algorithm below to obtain a set of new vectors $g^1, \dots, g^r \in \text{span}(b^1, \dots, b^d)$ with $r \leq d$. This is the part not formally justified in the paper by Bravyi et al [11]. The operators corresponding to these vectors will be linearly independent and commuting. We are thus looking for generators of a maximal Abelian subgroup of $\ker(E)$. Hence

$$\forall i, j \in \{1, \dots, r\} : g^i \times g^j = 0. \quad (24)$$

The algorithm for finding maximal abelian subgroups [21] is described as follows

1. Define $G = \langle \sigma(b^1), \dots, \sigma(b^d) \rangle$.
2. Now compute the center C of G and set $A := C$. Since A is abelian $A \subseteq C_G(A)$.
3. If $C_G(A) = A$ then A is a maximal abelian subgroup of G . Else, take an arbitrary $a \in C_G(A) \setminus A$. Define $A := \langle A \cup \{a\} \rangle$ and repeat step 3.

This procedure must halt as G is finitely generated and A keeps growing while still $A \subseteq G$ holds. As mentioned previously, it is known that there are maximally m generators in S [19]. Thus $S = \langle \tau_1, \tau_2, \dots, \tau_r \rangle$ where $r \leq m$.

Definition 6. Generator transform

Let $i \in 1, \dots, r$ and let τ_i be a generator of a symmetry group $S = \langle \tau_1, \dots, \tau_r \rangle$. Let τ_i anti-commute with a single

qubit operator σ_j^ρ , where $j \in 1, \dots, m$ and $\rho \in \{X, Y, Z\}$. Define new operators $\tilde{\tau}_k$ as

$$\tilde{\tau}_k = \begin{cases} \tau_k & \text{if } k = i \vee \tau_k \sigma_j^\rho = \sigma_j^\rho \tau_k \\ \tau_i \tau_k & \text{else} \end{cases} \quad (25)$$

These new operators are linearly independent, mutually commuting and generate S according to lemma 2. Furthermore, these new generators were transformed in such a way that $\tilde{\tau}_i$ anti-commutes with σ_j^ρ while all other $\tilde{\tau}_k$ commute with σ_j^ρ according to lemma 3.

For each $n \leq r$ the qubit reduction scheme requires a triple $\{A_n, B_n, V_n\}$ with $A_n = \{q(i)|i = 1, \dots, n \ i \neq j \rightarrow q(i) \neq q(j)\}$ a set of qubit-indices, $B_n = \{\rho(i)|i = 1, \dots, n \ \rho(i) \in \{X, Z\}\}$ a set of labels and $V_n = \{\nu_i|i = 1, \dots, r, \nu_i = \tau_1^{\alpha_1} \dots \tau_r^{\alpha_r}, \alpha_j \in \{0, 1\}\}$ a set of generators of S . This triple $\{A_n, B_n, V_n\}$ should satisfy the predicate

$$\begin{aligned} \Pi(\{A_n, B_n, V_n\}) : \quad & \forall i \in \{1, \dots, n\} \quad \forall j \in \{1, \dots, r\} \\ & \sigma_{q(i)}^{\rho(i)} \nu_j = (-1)^{\delta_{i,j}} \nu_j \sigma_{q(i)}^{\rho(i)} \end{aligned} \quad (26)$$

This requirement states that each qubit i should anti-commute only with generator i .

Lemma 4. Assume that $\Pi(\{A_n, B_n, V_n\})$ is satisfied. Let $\tau_{n+1} = P_1 P_2 \dots P_m$ where $P_i \in \{I, X, Y, Z\}$. There exists $j \in \{1, \dots, m\}$ such that $j \notin A_n$ and $P_j \neq I$.

Proof. Proof by contradiction. Assume that $\forall i \notin A_n : P_i = I$. Define the subgroup $S_l^n < P_m$ for $l \leq n$ as

$$\begin{aligned} S_l^n(\{A_n, B_n, V_n\}) = \\ \{ \sigma = P_1 P_2 \dots P_m \mid j \notin A_n \Rightarrow P_j = I, \forall i \in \{1, \dots, l\} \\ \sigma \tau_i = \tau_i \sigma, \forall w \in \{1, \dots, n\} \sigma \sigma_{q(w)}^{\rho(w)} = \sigma_{q(w)}^{\rho(w)} \sigma \} \end{aligned} \quad (27)$$

S_l^n is a subgroup of P_m as for every $l \leq n$ it is true that $I^{\otimes m} \in S_l$ and if $\sigma_a, \sigma_b \in S_l^n$ then $\sigma_a \sigma_b \in S_l^n$ (the inverses are of course trivially satisfied as every element is its own inverse). One can see that $\tau_{n+1} \in S_n^n$ by construction. From the condition $\forall w \in \{1, \dots, n\} \sigma \sigma_{q(w)}^{\rho(w)} = \sigma_{q(w)}^{\rho(w)} \sigma$ it follows that $|S_0^n| = 2^n$, as in every position in A_n only two operators out of $\{I, X, Y, Z\}$ are allowed. It is easy to see that S_{l+1}^n is a subgroup of S_l^n . Lagrange's theorem gives

$$|S_l^n| / |S_{l+1}^n| = |S_l^n / S_{l+1}^n| \in \mathbb{N}. \quad (28)$$

Note now that $\sigma_{q(l+1)}^{\rho(l+1)}$ is in S_l^n but not in S_{l+1}^n , since it must anti commute with τ_{l+1} . So, $|S_l^n| / |S_{l+1}^n| \geq 2$. Then $|S_0^n| / |S_n^n| \geq 2^n$. Therefore, $|S_n^n| = 1$ and thus $S_n^n = I$, the only group with one element. But by construction it was known that $\tau_{n+1} \in S_n^n$. Since I can not be a generator a contradiction is reached. Therefore, there must indeed exist a $j \in 1, \dots, m$ such that $j \notin A_n$ and $P_j \neq I$. \square

Theorem 5. Let $S = \langle \tau_1, \tau_2, \dots, \tau_r \rangle$ be a symmetry group. For every $n \leq r$ there exists a triple $\{A_n, B_n, V_n\}$ such that $\Pi(\{A_n, B_n, V_n\})$ is satisfied.

Proof. The proof is given by induction over n

Base case $n = 1$: consider $\tau_1 = P_1 P_2 \dots P_m$ where $P_i \in \{I, X, Y, Z\}$. Since τ_1 is a generator there must be a $P_j \neq I$. The first qubit index is chosen as $q(1) = j$ thus $A_1 = \{j\}$. The first label $\rho(1)$ is chosen as

$$\rho(1) = \begin{cases} X & \text{if } P_j \in \{Y, Z\} \\ Z & \text{if } P_j = X \end{cases} . \quad (29)$$

Thus, $B_1 = \{\rho(1)\}$. The generator transform from definition 6 can now be applied to the original generators of S to yield new generators $\tilde{\tau}_1, \dots, \tilde{\tau}_r$ which form V_1 . For ease of argumentation the new generators $\tilde{\tau}_1, \dots, \tilde{\tau}_r$ are renamed to τ_1, \dots, τ_r . Lemmas 2 and 3 now show that $\Pi(\{A_1, B_1, V_1\})$ is satisfied.

Induction cases $2 \leq n+1 \leq r$.

From lemma 4 there exists a $j \in 1, \dots, m$ such that $j \notin A_n$ and $P_j \neq I$. One can then choose $q(n+1) = j$ and

$$\rho(n+1) = \begin{cases} X & \text{if } P_j \in \{Y, Z\} \\ Z & \text{if } P_j = X \end{cases} . \quad (30)$$

The generator transform of definition 6 can be applied to the generators in V_n to yield new generators $\tilde{\tau}_1, \dots, \tilde{\tau}_r$. These generators form V_{n+1} . For ease of argumentation the new generators $\tilde{\tau}_1, \dots, \tilde{\tau}_r$ are renamed to τ_1, \dots, τ_r . $\Pi(\{A_{n+1}, B_{n+1}, V_{n+1}\})$ is now satisfied. \square

The existence and construction of a triple $\{A_r, B_r, V_r\}$ such that $\Pi(\{A_r, B_r, V_r\})$ is satisfied has been proven. Using the found triple $\{A_r, B_r, V_r\}$ one can define the operators U_i for $i = 1, \dots, r$, according to

$$U_i = \begin{cases} \frac{1}{2}(\sigma_{q(i)}^z + \tau_i)(\sigma_{q(i)}^z + \sigma_{q(i)}^x) & \text{if } \rho(i) = Z \\ \frac{1}{\sqrt{2}}(\sigma_{q(i)}^x + \tau_i) & \text{if } \rho(i) = X \end{cases} . \quad (31)$$

Below certain properties of the defined operators U_i are shown. Since Pauli matrices are Hermitian and unitary $\sigma_{q(i)}^{\rho(i)\dagger} = \sigma_{q(i)}^{\rho(i)}$, $\tau_i^\dagger = \tau_i$ and $\sigma_{q(i)}^{\rho(i)\dagger} \sigma_{q(i)}^{\rho(i)} = \tau_i^\dagger \tau_i = 1$. Using the implied commutation relations the following holds

$$\begin{aligned} U_i^\dagger U_i &= 1, & U_i \sigma_{q(i)}^x U_i^\dagger &= \tau_i, \\ U_i \sigma_{q(j)}^x U_i^\dagger &= \sigma_{q(j)}^x & \text{for } j &\neq i. \end{aligned} \quad (32)$$

This allows for the definition of the unitary operator $U = U_1 U_2 \dots U_r W$, where W is a permutation operator assigning qubit i to qubit $q(i)$. As a combination of unitary operators U is also unitary. The following relation follows from Eq. (32) $U \sigma_i^x U^\dagger = \tau_i$.

The transformed Hamiltonian H' can be defined as

$$H' = U^\dagger H U = \sum_k h_k U^\dagger \sigma^k U. \quad (33)$$

Define $\boldsymbol{\eta}^k = U^\dagger \sigma^k U$. Since $\forall i, j \in \{1, \dots, r\} [\sigma^i, \tau_j] = 0$ it must be that $\forall i, j \in \{1, \dots, r\} [\boldsymbol{\eta}^i, \sigma_j^x] = 0$. Therefore,

H' commutes with the last r qubits and thus shares eigenvectors with $\sigma_{m-(r-1)}^x, \dots, \sigma_m^x$. The last r qubits can thus be replaced by the X operator eigenvalues ± 1 .

This section justifies the \mathbb{Z}_2 symmetry qubit reduction scheme described by Bravyi et al. [11]. It shows a constructive proof of the scheme. Since it is constructive it should be feasible to implement this scheme in code and put it to use. Bravyi et al. have successfully applied this scheme in practice to Hamiltonians of small molecules such as H_2 and LiH to reduce the number of qubits necessary to describe the problem by 1. Because the focus of this paper is on the individual influences of the entangling methods and the classical optimization algorithms on the VQE algorithm, the reduction algorithm will not be used in the rest of the paper.

V. MULTI-QUBIT ENTANGLING METHODS

In the preparation of the trial state the entanglement operator U_{ent} ensures that the full Hilbert space of trial states can be reached. In this section multi-qubit entangling methods will be described, some gate based and others native to the physical qubit systems. In this work we study the Rydberg quantum computing system and its accompanying Rydberg interaction [22–25]. This section discusses the multi-qubit gates used in the simulations of Sec. VII.

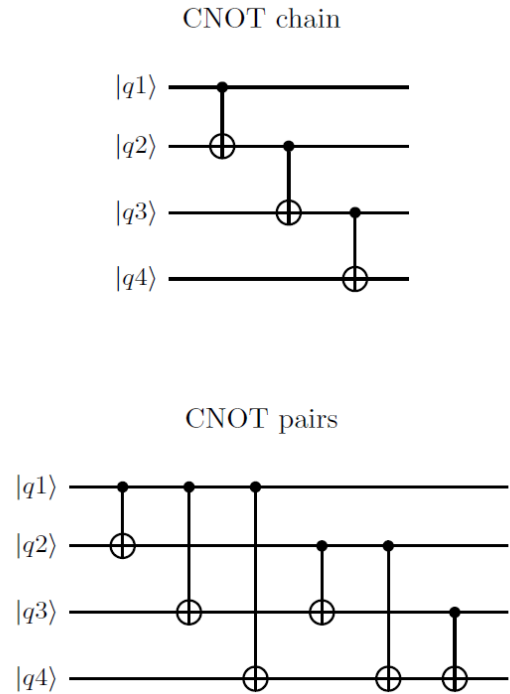


Figure 2: Quantum circuit diagram of the CNOT chain and pairs methods for $m = 4$.

The standard entanglement operator on two qubits is the CNOT gate. A way to entangle more than two qubits is by applying the CNOT gate in a chain or pairwise as is shown in Fig. 2. Since qubit operations are prone to errors it is desirable to use as few as possible to entangle all qubits. The chain and pair methods respectively use $m - 1$ and $m(m - 1)/2$ CNOT operations for m qubits. A way of reducing the number of operations is using entanglement interactions native to the qubit system. These interactions can entangle all qubits using only one operation and thus when considered as a single gate, are a more efficient way of entangling all qubits. Note that this efficiency only holds if the interaction is native to the qubit system, on other systems it might have to be constructed out of multiple other interactions.

A method proposed by Jaksch et al. [24] uses the Rydberg interaction. The Rydberg interaction is a long range interaction which, when one Rydberg atom (qubit) is in the excited state, blocks the others from going to that state. Thus it can function much like a CNOT gate. However, because of the long range interaction it also allows for a so called C_{m-1} NOT gate where one qubit controls all $m - 1$ other qubits, thus multi-qubit entanglement is achieved [26]. Another physical multi-qubit entangling method is the Krawtchouk chain [27]. The Krawtchouk chain is an entangling method based on the physical Hamiltonian of a Rydberg system which can be described as an interacting spin chain. Such a spin chain could consist of trapped ions or ultra-cold Rydberg atoms [27]. The Krawtchouk chain allows for the construction of the iPHASE $_m$ gate which mirrors the left and right side of a chain of qubits and multiplies it by a factor i (for instance $|01101\rangle \rightarrow i|10110\rangle$).

The Krawtchouk chain is an important method to construct other multi-qubit gates. Several studies have shown that the Krawtchouk Hamiltonian gate can be transformed, using only few one and two qubit rotations [28, 29]. Two examples of these transformations are the iSWAP $_2$ gate and the C_{m-1} NOT gate. The iSWAP $_2$ gate swaps the last two qubits and multiplies the state by a phase i . This swapping is controlled by the first $m - 2$ qubits. Figure 3 shows how the iPHASE $_m$ gate can be transformed into a C_{m-1} NOT gate.

VI. CLASSICAL OPTIMIZATION METHODS

Alongside the multi-qubit entangling method the classical optimization method plays an important role in the VQE algorithm. In this paper two contrasting methods which have been applied to VQE in previous research, are used. The local Simultaneous Perturbation Stochastic Approach (SPSA) [4] and the global DIviding RECTangles (DIRECT) method [30, 31].

SPSA starts with an input initial trial state $|\Psi(\vec{\theta}_0)\rangle$

with parameter vector $\vec{\theta}_0 \in [0, 2\pi]^{\otimes D}$. At every iteration step a small perturbation is done from the parameter vector towards a random direction to obtain two new parameter vectors $\vec{\theta}_{k,\pm} = \vec{\theta}_k + c_k \vec{\Delta}_k$. Here $\vec{\Delta}_k$ is a $(3d + 2)m$ -dimensional random vector that is generated by a sufficiently random generator and c_k is a yet to be defined constant. In this paper $\vec{\Delta}_k$ is determined by the Rademacher distribution where every vector entry has a probability of 1/2 of being either 1 or -1. The gradient \vec{g}_k is estimated by

$$\vec{g}_k = \frac{\langle \Psi(\vec{\theta}_{k,+}) | H | \Psi(\vec{\theta}_{k,+}) \rangle - \langle \Psi(\vec{\theta}_{k,-}) | H | \Psi(\vec{\theta}_{k,-}) \rangle}{2c_k} \vec{\Delta}_k. \quad (34)$$

To decrease the energy of the trial state at iteration k a step against the gradient direction is taken such that

$$\vec{\theta}_{k+1} = \vec{\theta}_k - a_k \vec{g}_k, \quad (35)$$

where a_k is a weight factor dependent on the iteration step. The factors a_k and c_k are given by

$$a_k = \frac{a}{k^A}, \quad c_k = \frac{c}{k^\Gamma}. \quad (36)$$

The constants A, c and Γ are best chosen to be 0.602, 0.01 and 0.101 respectively to guarantee the smoothest gradient descent [32]. The constant a is dependent on the gradient around the initial position and is defined according to

$$a = \frac{2\pi}{5} \frac{c}{\left\langle \left| \langle \Psi(\vec{\theta}_{1,+}) | H | \Psi(\vec{\theta}_{1,+}) \rangle - \langle \Psi(\vec{\theta}_{1,-}) | H | \Psi(\vec{\theta}_{1,-}) \rangle \right| \right\rangle_{\vec{\Delta}_1}}, \quad (37)$$

where $\langle \dots \rangle_{\vec{\Delta}_1}$ is the expectation value over the distribution of $\vec{\Delta}_1$. This is the way the average slope around the initialization vector is calculated in order to calibrate the optimization process. In practice this expectation value is approximated using Monte-Carlo methods. A step is taken in the direction opposite of the gradient in order to, conceivably, reach a lower energy value. The SPSA algorithm is a local optimization method.

The second method is the DIviding RECTangles (DIRECT) approach [33, 34]. The DIRECT algorithm attempts to find the global minimum of a Lipschitz continuous function $f : [0, 2\pi]^{\otimes D} \rightarrow \mathbb{R}$ defined on a convex hypercube $[0, 2\pi]^{\otimes D}$. A Lipschitz continuous function is one for which the following relation holds

$$\exists K > 0 \forall \vec{x}_1, \vec{x}_2 \in [0, 2\pi]^{\otimes D} : |f(\vec{x}_2) - f(\vec{x}_1)| \leq K \|\vec{x}_2 - \vec{x}_1\|_2. \quad (38)$$

Appendix A shows the proof that Hamiltonian evaluation, $f(\vec{x}) = \langle \vec{x} | H | \vec{x} \rangle$ is Lipschitz continuous.

A hypercube s is a cuboid subset of \mathbb{R}^D . Let S_k be the set of hypercubes that the original hypercube $[0, 2\pi]^{\otimes D}$

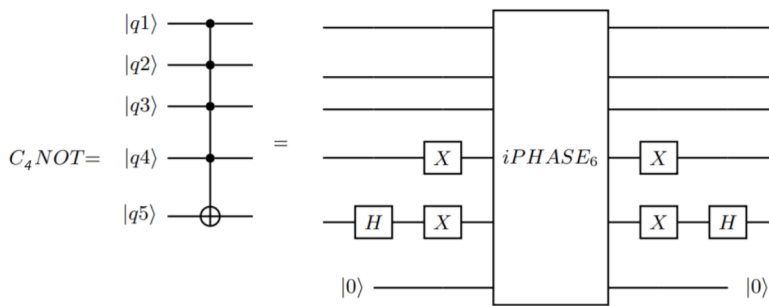


Figure 3: Quantum circuit diagram for the C_4NOT gate for $m = 5$ together with how it would be constructed from the $iPHASE_6$ gate which is created by the Krawtchouk Hamiltonian. Here H is a Hadamard gate.

was divided in after k iterations. Each of these hypercubes $s_j^k \in S_k$ has a center $c_j^k \in \mathbb{R}^D$ and a dimension $d_j^{k,max} \in \mathbb{R}$, the length of its largest side. $d_j^{k,i}$ will be the length of its side in direction i . In the DIRECT algorithm optimal hypercubes will be partitioned in the next iteration. A hypercube $s_i^k \in S_k$ is considered optimal if there exists a $\tilde{K} > 0$ such that

$$f(c_i^k) - \tilde{K}d_i^{k,max} \leq f(c_j^k) - \tilde{K}d_j^{k,max}, \quad j = 1, \dots, |S_k| \quad (39a)$$

$$f(c_j^k) - \tilde{K}d_j^{k,max} \leq f_{min}^k - \epsilon |f_{min}^k|, \quad (39b)$$

where $f_{min}^k = \min_j \{f(c_j^k)\}$, and ϵ is a small positive constant. In this research $\epsilon = 10^{-20}$. An optimal hypercube s_j^k with center c_j^k will be partitioned into smaller hypercubes. For a subdivision first a set of new sample points in the hypercube is defined as

$$A_j^{k+1} = \{c_j^k\} \cup \{c_j^k \pm \frac{1}{3}d_j^{k,max}\vec{e}_i \mid i : d_j^{k,i} = d_j^{k,max}\}, \quad (40)$$

where \vec{e}_i is the unit vector in direction i . So partitions are only done in the direction where the hypercube has the largest length. Since hypercubes can not overlap the order of partitioning the different dimensions matters. The first division will be done in the direction of the sample point with the lowest function value. The second division with the second lowest function value, etc., see Fig. 4. This way the hypercube with the lowest function value will become the largest of the partition. Larger hypercubes are more likely to satisfy Eq. (39a) and thus to be partitioned again in the next iteration. The partitioning will result in a new set of hypercubes S_{k+1} which can be tested for optimality. The center of an optimal hypercube S_k will again be the center of a hypercube in S_{k+1} . The set of centers for which the energy is tested thus grows at each iterations step. The hypercubes will eventually converge to the optimal ground state. The DIRECT algorithm is a global optimization method.

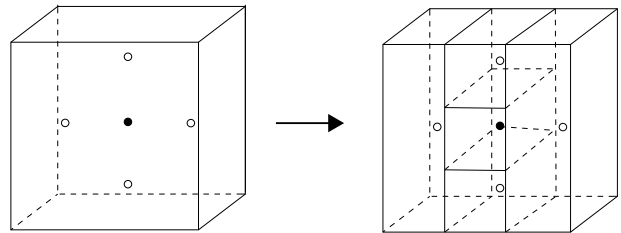


Figure 4: Iteration step of the DIRECT algorithm for a three dimensional ($D = 3$) hypercube. Divisions are only done along the dimensions with the largest length.

VII. RESULTS

This section will investigate the behaviour of the discussed multi-qubit entangling methods and classical optimization algorithms on VQE problems of H_2 and LiH. For both H_2 and LiH the simulations are run over interatomic distances from $0.392a_0$ to $0.931a_0$ and $0.72a_0$ to $1.71a_0$ respectively, in steps of $0.049a_0$ and $0.09a_0$ respectively, with a_0 the Bohr-radius, and the resulting errors with the Hartree-Fock energy are averaged. For problems with a small number of qubits, such as H_2 with the $1s$ orbital (2 qubits) and LiH with the $1s$, $2s$ and $2p_x$ orbitals (6 qubits), the dimension of the state space is limited. In other, more complex, problems one might want to start off with a certain ansatz initial state to see if the VQE algorithm is able to improve upon this. For consistency, we choose the initial state to be the vacuum state $|\psi_{init}\rangle = |0\dots 0\rangle$ in this work. When the errors in ground state energies of two algorithms are compared, the algorithm with the lower error is said to outperform the other.

Table I: Errors in Hartree for exact ground energy for 10^3 iteration SPSA simulations comparing the entangling methods for H_2 and LiH at depths 1 to 6.

Molecule	Entangler	d=1	d=2	d=3	d=4	d=5	d=6
H_2	CNOT chain	4.8×10^{-4}	1.8×10^{-4}	1.9×10^{-4}	3.2×10^{-4}	4.1×10^{-4}	5.9×10^{-4}
	C_k NOT	2.1×10^{-4}	1.5×10^{-4}	1.8×10^{-4}	3.1×10^{-4}	3.2×10^{-4}	6.0×10^{-4}
	iPHASE $_m$	1.9×10^{-2}	1.9×10^{-2}	1.9×10^{-2}	1.9×10^{-2}	1.9×10^{-2}	1.9×10^{-2}
	iSWAP $_2$	1.9×10^{-2}	1.9×10^{-2}	1.9×10^{-2}	1.9×10^{-2}	1.9×10^{-2}	1.9×10^{-2}
LiH	CNOT chain	1.0×10^{-1}	3.1×10^{-2}	1.9×10^{-2}	1.6×10^{-3}	1.5×10^{-2}	1.5×10^{-3}
	C_k NOT	1.2×10^{-1}	2.6×10^{-2}	1.8×10^{-2}	5.0×10^{-3}	1.5×10^{-2}	6.6×10^{-3}
	iPHASE $_m$	4.2×10^{-3}	2.9×10^{-3}	1.9×10^{-3}	1.4×10^{-3}	1.4×10^{-2}	1.4×10^{-2}
	iSWAP $_2$	2.4×10^{-2}	2.1×10^{-3}	4.3×10^{-2}	1.9×10^{-3}	2.3×10^{-3}	2.0×10^{-3}

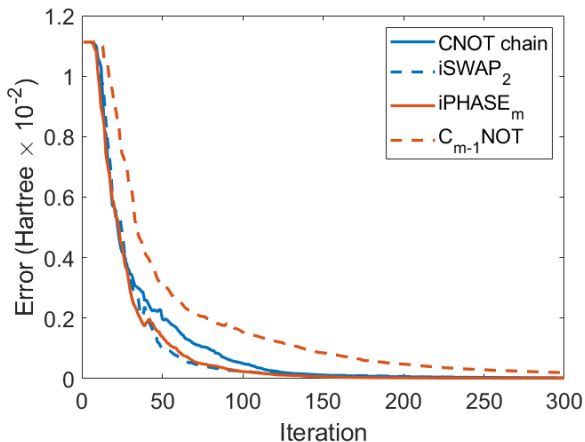


Figure 5: A typical energy error from Hartree-Fock energy versus iteration number for the entangling methods iPHASE $_m$, C_{m-1} NOT and iSWAP $_2$ compared to the standard CNOT chain. The VQE problem describes a LiH molecule (6 qubits) at depth 6, with classical optimization SPSA.

To investigate the dependence of the VQE results on entangling methods for fixed (SPSA) classical optimization algorithm, we first analyze the ground state of LiH for $d=6$ using the iPHASE $_m$, iSWAP $_2$, C_{m-1} NOT and CNOT chain gates (discussed in Sec. V) as shown in Fig. 5. Here it is found that methods based on the Rydberg interaction are able to perform comparably to the CNOT chain method (see Sec. V.) Repeating these simulations for LiH and H_2 as a function of depth d we get the resulting errors given in Table I. These results seem to further support that Rydberg interaction entangling methods can compare to the CNOT chain method in terms of performance. The optimal entangling method however appears to be problem dependent. The entanglement depth d can open up the search space for the trial state allowing lower errors to be reached as can be seen in many of the LiH cases of Table I. However, it

can also overcomplicate the problem by introducing too many rotation parameters and thus take more iterations to converge as we see in H_2 cases. Another interesting fact is seen for iPHASE $_m$ and iSWAP $_2$ at H_2 , where the entangling method is not fit for the problem regardless of the depth, and an error as low as the CNOT chain and C_k NOT errors is not reached.

Next, we analyze the ground state energies of LiH and H_2 as a function of the classical algorithm optimization algorithm (SPSA or DIRECT, see Sec. VI) for both CNOT and iPHASE $_m$ to check additional dependence on the entangling method as shown in Fig. 6. We summarize the results of the VQE calculations for a specific molecule in terms of the depth and dimensionality. The considered VQE problems are Hamiltonians of H_2 at depths 1 to 8 and of LiH at depths d 1 to 4. This results in dimensions ranging from 10 to 84. In order to compare SPSA and DIRECT, ground state energy computation is iterated for 10^4 and 25 times respectively. These numbers were chosen because for considered ground state calculations, the errors of the SPSA and DIRECT methods do not decrease by more than 10^{-4} Hartree in the last 10^3 and 5 iterations, respectively.

From Fig. 6 it can be seen that for H_2 ((a) and (c)) SPSA performs better, while for LiH ((b) and (d)) the DIRECT algorithm performs better. However, the DIRECT algorithm performs better on average at higher depths regardless of the entangling method, although this holds to a lesser degree for the iPHASE $_m$ results in Fig. 6 (c). The DIRECT algorithm also performs better for larger numbers of qubits m , indicating that dimensionality D differentiates significantly the performance of classical search algorithms.

Having highlighted the role of dimensionality D , we return finally to discuss the relatively lesser dependence in Fig. 6 on the chosen entangling method. Comparing the results for CNOT pairs (Figs. 6 (a) and (b)) and

iPHASE_m (Figs. 6 (c) and (d)) it is seen that the errors reached in the same problems differ per entangling method. For instance at H₂, for CNOT pairs at high depths DIRECT gives less error than SPSA while for iPHASE_m SPSA still performs better. The preference of one algorithm over another seems thus to be dependent on the entangling method. It is remarkable that a different entanglement scheme changes the preference for a certain classical optimization algorithm as they are implemented in distinct parts of the algorithm. We speculate that this has to do with the “scrambling” nature of a given entangling method in a given search space. In a fully scrambled search space a gradient is meaningless. For this reason the SPSA algorithm might perform worse

compared to the DIRECT algorithm and thus a local search becomes expensive in terms of computing power. This would explain, for instance, why at high dimensionality the CNOT pairs method performs better when combined with a global rather than local search method (Figs. 6 (a) and (b)) and vice versa for the iPHASE_m method (Figs. 6 (c) and (d)). Ultimately this requires a way to measure the amount of scrambling that an entangling method performs on a given search space. Although no such measure exists presently for operators, it would be analogous to how the entropy of a state provides a measure of its degree of entanglement. The derivation of such a scrambling measure for operators on a given search space remains however the subject of future studies.

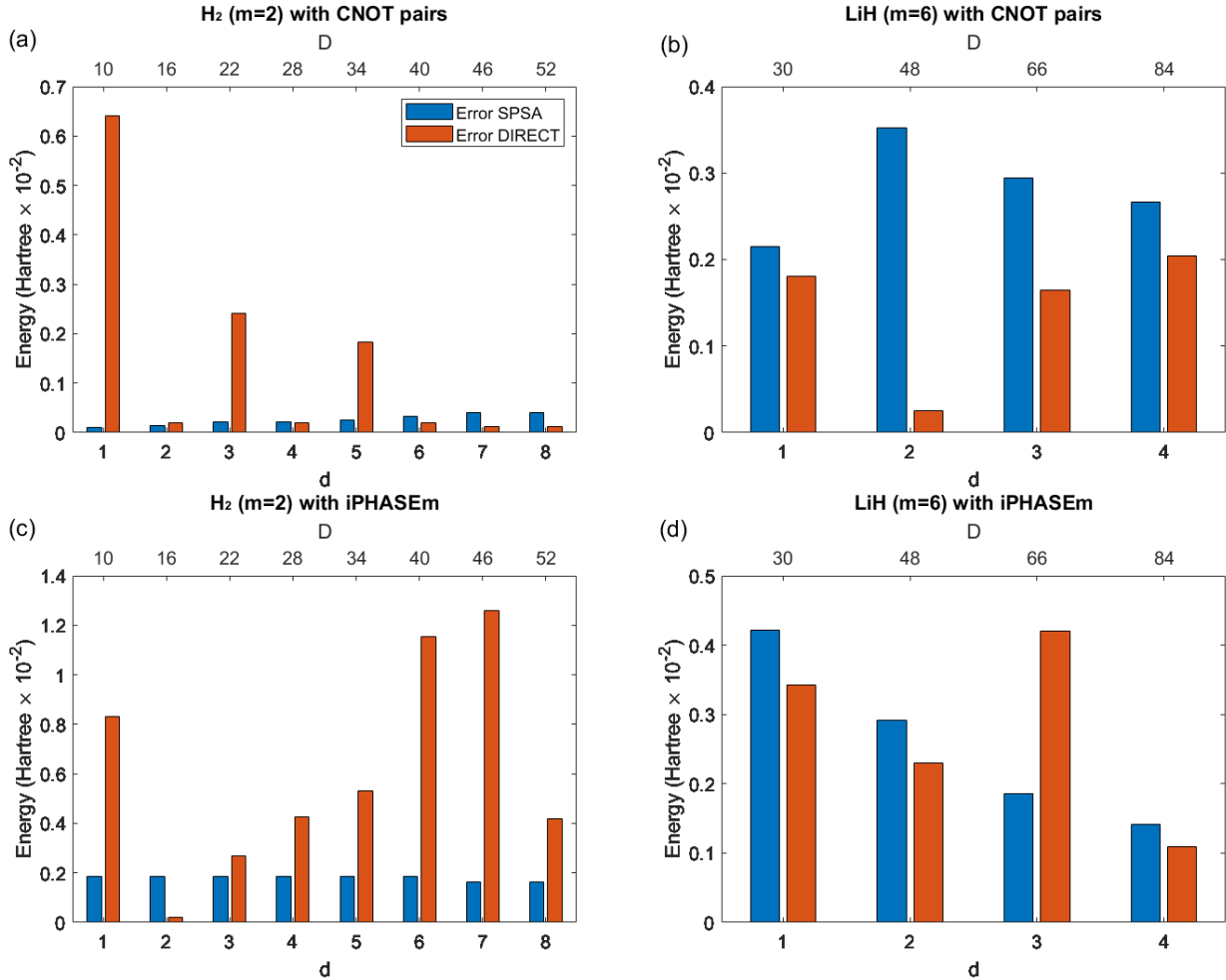


Figure 6: Error with the Hartree-Fock energy in the reached ground state energy for VQE problems with different depth, optimized with the SPSA and DIRECT algorithms. The entangling methods used are CNOT pairs ((a) and (b)) and iPHASE_m ((c) and (d)). Used VQE problems are the Hamiltonians of H₂ (2 qubits, (a) and (c)) at depths 1 to 8 and LiH (6 qubits, (b) and (d)) at depths 1 to 4. Thus, the problems have dimensions ranging from $D = 10$ to $D = 84$. SPSA was run with 10^4 iterations and DIRECT was run with 25 iterations. These were empirically chosen as was found that the next 10^3 and 5 iterations respectively yielded less than 10^{-4} Hartree improvement on the ground state energy.

VIII. CONCLUSION

In this paper, we have focused on optimizing the VQE algorithm for quantum chemistry. First, a constructive proof was shown for a qubit reduction scheme using the \mathbb{Z}_2 symmetries of Hamiltonians, based on the paper by Bravyi et al. [11], which effectively reduces the total number of qubit operations. Then we examined the performance of the VQE algorithm as a function of the entangling method choice, depth, and classical optimization algorithm. Entangling methods based on both physical interactions, such as iPHASE_m and C_{m-1}NOT gave a convergence that was comparable to that of the CNOT chain method in the tested small molecule VQE problems. From the results, although the entangling method and classical optimization method appear to interact in a complex, problem dependent manner, we found a general trend that global search algorithms perform better than local algorithms at higher dimensionality.

To further investigate links between dimensionality, entangling method, and classical optimization algorithm, VQE calculations should be performed for more complicated molecular systems. Furthermore, our findings motivate future studies on the scrambling powers of entanglement operators on a search space and how this might be encapsulated quite generally by some measure. Such a concept would be foundational towards the design of efficient VQE circuitry.

Data available on request from the authors.

ACKNOWLEDGEMENTS

We thank Jasper Postema and Gijs Groeneveld for discussions. This research is financially supported by the Netherlands Organisation for Scientific Research (NWO) under Grant No. 680-47-623 and 680.92.18.05.

-
- [1] J. Preskill, *Quantum* **2**, 79 (2018), ISSN 2521-327X, URL <http://dx.doi.org/10.22331/q-2018-08-06-79>.
- [2] F. Arute, K. Arya, R. Babbush, D. Bacon, J. C. Bardin, R. Barends, R. Biswas, S. Boixo, F. G. S. L. Brandao, D. A. Buell, et al., *Nature* **574**, 505 (2019), ISSN 1476-4687, URL <https://doi.org/10.1038/s41586-019-1666-5>.
- [3] A. Peruzzo, J. McClean, P. Shadbolt, M.-H. Yung, X.-Q. Zhou, P. J. Love, A. Aspuru-Guzik, and J. L. O'Brien, *Nature Communications* **5**, 4213 (2014), 1304.3061.
- [4] A. Kandala, A. Mezzacapo, K. Temme, M. Takita, M. Brink, J. M. Chow, and J. M. Gambetta, *Nature (London)* **549**, 242 (2017), 1704.05018.
- [5] C. Hempel, C. Maier, J. Romero, J. McClean, T. Monz, H. Shen, P. Jurcevic, B. P. Lanyon, P. Love, R. Babbush, et al., *Phys. Rev. X* **8**, 031022 (2018), URL <https://link.aps.org/doi/10.1103/PhysRevX.8.031022>.
- [6] A. Aspuru-Guzik and P. Walther, *Nature Physics* **8** (2012).
- [7] T. O'Brien, P. Rožek, and A. Akhmerov, *Physical review letters* **120**, 220504 (2018).
- [8] F. and Arute, K. Arya, R. Babbush, D. Bacon, J. C. Bardin, R. Barends, S. Boixo, M. Broughton, B. B. Buckley, D. A. Buell, et al., *Science* **369**, 1084 (2020), ISSN 0036-8075, URL <https://science.sciencemag.org/content/369/6507/1084>.
- [9] I. B. Jonathan Koller, William Bender, ed., *Exploring Quantum Approximations of Traveling Salesman* (Seidenberg School of CSIS, Pace University, Pleasantville, New York 10570, 2019).
- [10] IBM-Research-Editorial-Staff, *How to measure a molecule's energy using a quantum computer* (2017), URL <https://www.ibm.com/blogs/research/2017/09/quantum-molecule/>.
- [11] S. Bravyi, J. M. Gambetta, A. Mezzacapo, and K. Temme, arXiv e-prints arXiv:1701.08213 (2017), 1701.08213.
- [12] S. McArdle, S. Endo, A. Aspuru-Guzik, S. Benjamin, and X. Yuan, arXiv e-prints arXiv:1808.10402 (2018), 1808.10402.
- [13] Note1, hartree units are also known as atomic units where the reduced constant of Plank \hbar , the electron mass m_e , the elementary charge e and Coulomb's constant $k_e = 1/4\pi\epsilon_0$ are equal to unity. In this system the Bohr radius $a_0 = 4\pi\epsilon_0\hbar^2/m_e e^2$ is also equal to unity. Hartree units are often used in molecular level calculations [15].
- [14] J. R. McClean, J. Romero, R. Babbush, and A. Aspuru-Guzik, *New Journal of Physics* **18**, 023023 (2016), 1509.04279.
- [15] D. McQuarrie, *Quantum Chemistry*, v. 1 (University Science Books, 2008), ISBN 9781891389504, URL <https://books.google.nl/books?id=zzxLTI1jQB4C>.
- [16] Note2, in this work the integrals in Eq. 3 and Eq. 4 have been determined using quantum computation library *OpenFermion* [35].
- [17] P. Coleman, *Simple examples of second quantization* (Cambridge University Press, 2015), p. 71–94.
- [18] T. A. Brun, *Quantum error correction* (2019), 1910.03672.
- [19] K. Fujii, *Quantum Computation with Topological Codes*, vol. 1 of 8 (Springer Singapore, 2015), 1st ed., ISBN 978-981-287-995-0.
- [20] D. Gottesman, Ph.D. thesis, California Institute of Technology (1997).
- [21] A. G. Kurosh, *Theory of groups* (Amer Mathematical Society, 2014).
- [22] M. Saffman, *Journal of Physics B: Atomic, Molecular and Optical Physics* **49**, 202001 (2016), URL <https://doi.org/10.1088/1361-6455/49/20/202001>.

- [org/10.1088/0953-4075/49/20/202001](https://doi.org/10.1088/0953-4075/49/20/202001).
- [23] C. S. Adams, J. D. Pritchard, and J. P. Shaffer, *Journal of Physics B: Atomic, Molecular and Optical Physics* **53**, 012002 (2019), URL <https://doi.org/10.1088/1361-6455/ab52ef>.
- [24] D. Jaksch, J. I. Cirac, P. Zoller, S. L. Rolston, R. Côté, and M. D. Lukin, *Phys. Rev. Lett.* **85**, 2208 (2000), [quant-ph/0004038](https://arxiv.org/abs/quant-ph/0004038).
- [25] M. Morgado and S. Whitlock, *Quantum simulation and computing with rydberg-interacting qubits* (2020), 2011.03031.
- [26] T. G. Walker and M. Saffman, in *Advances in Atomic, Molecular, and Optical Physics*, edited by P. Berman, E. Arimondo, and C. Lin (Academic Press, 2012), vol. 61 of *Advances In Atomic, Molecular, and Optical Physics*, pp. 81 – 115, URL <http://www.sciencedirect.com/science/article/pii/B9780123964823000028>.
- [27] H. Labuhn, D. Barredo, S. Ravets, S. de Léséleuc, T. Macrì, T. Lahaye, and A. Browaeys, *Nature (London)* **534**, 667 (2016), 1509.04543.
- [28] K. Groenland and K. Schoutens, *Phys. Rev. A* **97**, 042321 (2018), 1707.05144.
- [29] S. R. Clark, C. Moura Alves, and D. Jaksch, *New Journal of Physics* **7**, 124 (2005), [quant-ph/0406150](https://arxiv.org/abs/quant-ph/0406150).
- [30] C. Kokail, C. Maier, R. van Bijnen, T. Brydges, M. K. Joshi, P. Jurcevic, C. A. Muschik, P. Silvi, R. Blatt, C. F. Roos, et al., *Nature* **569**, 355–360 (2019), ISSN 1476-4687, URL <http://dx.doi.org/10.1038/s41586-019-1177-4>.
- [31] C. Kokail, C. Maier, R. van Bijnen, T. Brydges, M. K. Joshi, P. Jurcevic, C. A. Muschik, P. Silvi, R. Blatt, C. F. Roos, et al., *Nature* **569**, 355–360 (2019), ISSN 1476-4687, URL <http://dx.doi.org/10.1038/s41586-019-1177-4>.
- [32] J. C. Spall, *IEEE Transactions on Aerospace and Electronic Systems* **34**, 817 (1998), ISSN 0018-9251.
- [33] P. J. Nicholas, in *A Dividing Rectangles Algorithm for Stochastic Simulation Optimization* (2015).
- [34] M. Piao Tan and C. A. Floudas, *Determining the optimal number of clusters* *Determining the Optimal Number of Clusters* (Springer US, Boston, MA, 2009), pp. 687–694, ISBN 978-0-387-74759-0, URL https://doi.org/10.1007/978-0-387-74759-0_123.
- [35] J. R. McClean, K. J. Sung, I. D. Kivlichan, Y. Cao, C. Dai, E. Schuyler Fried, C. Gidney, B. Gimby, P. Gokhale, T. Häner, et al., *arXiv e-prints* [arXiv:1710.07629](https://arxiv.org/abs/1710.07629) (2017), 1710.07629.

Appendix A: Proof Lipschitz continuity of Hamiltonian evaluation

Theorem 6. Let H be a $2^m \times 2^m$ Hamiltonian operator. The function $f(\vec{x}) = \langle \Psi(\vec{x}) | H | \Psi(\vec{x}) \rangle$, where \vec{x} describes a rotation of an ansatz $|\phi\rangle$, is Lipschitz continuous.

Proof.

$$\begin{aligned}
|f(\vec{x}_2) - f(\vec{x}_1)| &= |\langle \Psi(\vec{x}_2) | H | \Psi(\vec{x}_2) \rangle - \langle \Psi(\vec{x}_1) | H | \Psi(\vec{x}_1) \rangle| \\
&= |\langle \Psi(\vec{x}_2) | H | \Psi(\vec{x}_2) \rangle - \langle \Psi(\vec{x}_2) | H | \Psi(\vec{x}_1) \rangle + \langle \Psi(\vec{x}_2) | H | \Psi(\vec{x}_1) \rangle - \langle \Psi(\vec{x}_1) | H | \Psi(\vec{x}_1) \rangle| \\
&= |\langle \Psi(\vec{x}_2) | H | \Psi(\vec{x}_2) - \Psi(\vec{x}_1) \rangle + \langle \Psi(\vec{x}_2) - \Psi(\vec{x}_1) | H | \Psi(\vec{x}_1) \rangle| \\
&\leq |\langle \Psi(\vec{x}_2) | H | \Psi(\vec{x}_2) - \Psi(\vec{x}_1) \rangle| + |\langle \Psi(\vec{x}_2) - \Psi(\vec{x}_1) | H | \Psi(\vec{x}_1) \rangle| \\
&\leq \|\Psi(\vec{x}_2)\|_2 \cdot \|H\|_2 \cdot \|\Psi(\vec{x}_2) - \Psi(\vec{x}_1)\|_2 + \|\Psi(\vec{x}_1)\|_2 \cdot \|H\|_2 \cdot \|\Psi(\vec{x}_2) - \Psi(\vec{x}_1)\|_2 \\
&= 2\|H\|_2 \cdot \|\Psi(\vec{x}_2) - \Psi(\vec{x}_1)\|_2.
\end{aligned} \tag{A1}$$

Now the trial state $|\Psi(\vec{x})\rangle$ is generated by rotations on the ansatz $|\phi\rangle$, so $|\Psi(\vec{x})\rangle = R(\vec{x})|\phi\rangle$. All of the rotation matrices as defined in Eq. (14) have elements which are continuously differentiable. Since the product of continuously differentiable functions is still continuously differentiable, it is true that $R(\vec{x})$ can be written as

$$R(\vec{x}) = \begin{pmatrix} f_{1,1}(\vec{x}) & f_{1,2}(\vec{x}) & \cdots & f_{1,2^m}(\vec{x}) \\ f_{2,1}(\vec{x}) & f_{2,2}(\vec{x}) & \cdots & f_{2,2^m}(\vec{x}) \\ \vdots & \vdots & \ddots & \vdots \\ f_{2^m,1}(\vec{x}) & f_{2^m,2}(\vec{x}) & \cdots & f_{2^m,2^m}(\vec{x}) \end{pmatrix}, \tag{A2}$$

where $f_{i,j}(\vec{x})$ are all continuously differentiable functions. Since all continuously differentiable functions are Lipschitz continuous

$$\forall i, j \exists a_{i,j} > 0 : |f_{i,j}(\vec{x}_2) - f_{i,j}(\vec{x}_1)| \leq a_{i,j} \|\vec{x}_2 - \vec{x}_1\|_2. \tag{A3}$$

Then it is true that

$$\begin{aligned}
\|\Psi(\vec{x}_2) - \Psi(\vec{x}_1)\|_2 &= \|R(\vec{x}_2)|00\dots 0\rangle - R(\vec{x}_1)|00\dots 0\rangle\| \leq \|R(\vec{x}_2) - R(\vec{x}_1)\|_2 \\
&= \left\| \begin{pmatrix} f_{1,1}(\vec{x}_2) & f_{1,2}(\vec{x}_2) & \cdots & f_{1,2^m}(\vec{x}_2) \\ f_{2,1}(\vec{x}_2) & f_{2,2}(\vec{x}_2) & \cdots & f_{2,2^m}(\vec{x}_2) \\ \vdots & \vdots & \ddots & \vdots \\ f_{2^m,1}(\vec{x}_2) & f_{2^m,2}(\vec{x}_2) & \cdots & f_{2^m,2^m}(\vec{x}_2) \end{pmatrix} - \begin{pmatrix} f_{1,1}(\vec{x}_1) & f_{1,2}(\vec{x}_1) & \cdots & f_{1,2^m}(\vec{x}_1) \\ f_{2,1}(\vec{x}_1) & f_{2,2}(\vec{x}_1) & \cdots & f_{2,2^m}(\vec{x}_1) \\ \vdots & \vdots & \ddots & \vdots \\ f_{2^m,1}(\vec{x}_1) & f_{2^m,2}(\vec{x}_1) & \cdots & f_{2^m,2^m}(\vec{x}_1) \end{pmatrix} \right\|_2 \\
&\leq \sum_{i=1}^{2^m} \sum_{j=1}^{2^m} \|f_{i,j}(\vec{x}_2) - f_{i,j}(\vec{x}_1)\|_2 \leq \sum_{i=1}^{2^m} \sum_{j=1}^{2^m} a_{i,j} \|\vec{x}_2 - \vec{x}_1\|_2 \leq K \|\vec{x}_2 - \vec{x}_1\|_2,
\end{aligned} \tag{A4}$$

where $K = \sum_{i=1}^{2^m} \sum_{j=1}^{2^m} a_{i,j} \in \mathbb{R}_+$. Thus $|f(\vec{x}_2) - f(\vec{x}_1)| \leq 2 \cdot \|H\|_2 \cdot K \cdot \|\vec{x}_2 - \vec{x}_1\|_2$ and thus f is Lipschitz continuous. \square