On the order problem in construction of unitary operators for the Variational Quantum Eigensolver

Artur F. Izmaylov,[1,2,*] Manuel Díaz-Tinoco,[1,2] and Robert A. Lang[1,2]

[1] *Department of Physical and Environmental Sciences,*
University of Toronto Scarborough, Ontario, M1C 1A4, Canada
[2] *Chemical Physics Theory Group, Department of Chemistry,*
University of Toronto, Toronto, Ontario, M5S 3H6, Canada

One of the main challenges in the Variational Quantum Eigensolver (VQE) framework is construction of the unitary transformation. The dimensionality of the space for unitary rotations of N qubits is $4^N - 1$, which makes the choice of a polynomial subset of generators exponentially difficult process. Moreover, due to non-commutativity of generators, the order in which they are used strongly affects results. Choosing the optimal order in a particular subset of generators requires testing the factorial number of combinations. We propose an approach based on the Lie algebra - Lie group connection and corresponding closure relations that systematically eliminates the order problem.

I. INTRODUCTION

One of the most promising applications of near-term universal gate quantum computation is solving the electronic structure problem. Currently, the most feasible route to this task is through the variational quantum eigensolver (VQE),[1] which is a hybrid technique involving an iterative minimization of the electronic expectation value

$$E = \min_{\boldsymbol{\tau}} \langle \bar{0} | \hat{U}^\dagger(\boldsymbol{\tau}) \hat{H} \hat{U}(\boldsymbol{\tau}) | \bar{0} \rangle, \tag{1}$$

involving quantum and classical computers. First, a classical computer suggests a trial unitary transformation $\hat{U}(\boldsymbol{\tau})$ that is encoded on a quantum computer as a circuit operating on the initial state of N qubits $|\bar{0}\rangle \equiv |0\rangle^{\otimes N}$. At the end of this circuit, one measures the expectation value of the qubit-space Hamiltonian $\hat{H}$, which is iso-spectral to the electronic Hamiltonian in the second quantized form. The measured electronic energy is provided to the classical computer that generates a next guess for the parameterized unitary transformation $\hat{U}(\boldsymbol{\tau})$. These cycles serve to minimize the energy expectation value and to approach the true electronic ground state energy.

Generators of N-qubit unitary operations up to a global phase correspond to $4^N - 1$ basis vectors of the $su(2^N)$ Lie algebra, N-qubit Pauli products

$$\hat{P} = \bigotimes_{j=1}^{N} \hat{\sigma}_j, \tag{2}$$

where $\hat{\sigma}_j$'s are the Pauli operators $\{\hat{x}_j, \hat{y}_j, \hat{z}_j\}$ or the 2×2 identity $\hat{1}_j$ for the jth qubit. Thus, any element of the corresponding unitary group $SU(2^N)$ can be presented as

$$\hat{U}(\boldsymbol{\tau}) = \prod_{k=1}^{M} e^{i \tau_k \hat{P}_k}, \tag{3}$$

where $M \leq 4^N - 1$. Generators $\hat{P}_k$ do not generally commute, therefore the individual exponents in Eq. (3) do not commute either. Yet, since all generators have been included, their order is not important. Any order will be able to represent any element of the $SU(2^N)$ Lie group. Only values of τ_k's are changing in representing a particular $SU(2^N)$ element with different orders in Eq. (3).

The representation in Eq. (3) is not a conventional representation of a Lie group, $\mathcal{G}$, as an exponential map of the corresponding Lie algebra, g, with generators $\{g_k\}$,

$$G = e^{\sum_k C_k g_k}, \ G \in \mathcal{G}, \tag{4}$$

where C_k are real or complex numbers.[2] We can arrive to Eq. (3) considering the following additional properties of $SU(2^N)$ and the exponential map. $SU(2^N)$ is a compact Lie group and the corresponding algebra $su(2^N)$ is compact as well. For compact Lie groups the exponential mapping of the Lie algebra is surjective.[3] Thus, using the conventional exponential mapping between $su(2^N)$ and $SU(2^N)$ one can generate universal covering group which is simply connected (identical to $SU(2^N)$). For simply connected groups, one can use globally analytic Baker-Campbell-Hausdorff reparametrizations of a single exponent (Eq. (4)) to a disentangled form of multiple exponents as in Eq. (3).[4]

In practice, one cannot use M that scales exponentially with N, this would defeat the purpose of involving quantum computers for efficient representation of $\hat{U}$.[5] Thus, one needs efficient heuristics to make $\hat{P}_k$ selection so that M scales polynomially with N. Two main polynomial heuristics have been suggested for $\hat{U}$: 1) methods based on the fermionic unitary forms restricted by the orbital excitation level and then transformed to the qubit space[1,6–8] (e.g. unitary coupled cluster singles and doubles (UCCSD)[1,7,9,10]) and 2) qubit-space techniques, where $\hat{P}_k$'s are selected based on the energy gradients with respect to τ_k's.[11–13]Methods in both categories provide particular subsets of $\hat{P}_k$'s in Eq. (3) and thus lead to the order problem: different orders of exponents in Eq. (3) are not equivalent in their ability to lower the ex-

pectation values via optimizing corresponding τ_k's. Recently, several works have shown how large variations can be from changing the order.[8,14,15]

In this work we provide an approach to remove the order problem by using algebraic closure of a $\hat{P}_k$ set into a Lie sub-algebra of $su(2^N)$ by considering all possible commutators within the set. Considering Lie algebra - Lie group connection through the exponential map, the sub-algebra closure guarantees generation of a closed sub-group of $SU(2^N)$ where the order of exponents is not essential. Thus, even in the absence of commutativity, the closed algebraic structure to be introduced is sufficient for the order problem removing.

II. THEORY

A. Lie algebraic consideration

We assume that a set of $\hat{P}_k$'s is selected based on some energy lowering heuristic. Formally, this set $\mathcal{S} = \{\hat{P}_k\}_{k=1}^M$ forms a Lie sub-algebra of M generators if for any pair $\hat{P}_i$ and $\hat{P}_j$

$$[\hat{P}_i, \hat{P}_j] = \sum_k c_{ij}^{(k)} \hat{P}_k, \ \ \hat{P}_i, \hat{P}_j, \hat{P}_k \in \mathcal{S} \quad (5)$$

where $c_{ij}^{(k)}$ are so-called structural constants.[16] Then the exponential map of $\mathcal{S}$ produces a Lie sub-group $\mathcal{G}_S$ with elements that can be given by Eq. (3). The order of $\hat{P}_k$'s forming the $\mathcal{S}$ sub-algebra does not matter because any order of their exponential products can represent points of a sub-manifold corresponding to $\mathcal{G}_S$. Here, again, one of the crucial elements is compactness of both $\mathcal{S}$ and $\mathcal{G}_S$.

If a single sub-algebra will not be sufficient to provide enough elements to lower the energy, one can use a set of sub-algebras, $\{\mathcal{S}_j\}$ to construct $\hat{U} = \hat{U}_1 \times ... \times \hat{U}_n$, where

$$\hat{U}_j(\boldsymbol{\tau}_j) = \prod_{k=1}^{M_j} e^{i\tau_k^{(j)} \hat{P}_k^{(j)}}, \ \ \hat{P}_k^{(j)} \in \mathcal{S}_j. \quad (6)$$

In this case, the order of exponents within each $\hat{U}_j$ will not affect the minimum of the energy expectation value, but the order of $\hat{U}_j$'s in $\hat{U}$ will. This provides a simple prescription of how to remove the order problem for any neighbouring pairs $\hat{U}_j$ and $\hat{U}_{j+1}$: one needs to obtain the closure of corresponding sub-algebras $\mathcal{S}_j$ and $\mathcal{S}_{j+1}$, $\mathcal{S}_{j,j+1}^{(c)} = [\mathcal{S}_j, \mathcal{S}_{j+1}]$, which is a sub-algebra that contains all elements of $\mathcal{S}_j$ and $\mathcal{S}_{j+1}$ as well as all possible commutators. Clearly, the number of generators of $\mathcal{S}_{j,j+1}^{(c)}$ will not be smaller than or equal to the sum of those for $\mathcal{S}_j$ and $\mathcal{S}_{j+1}$. $\mathcal{S}_{j,j+1}^{(c)}$ elements can be exponentiated to create $\hat{U}_{j,j+1}$, which will be an order invariant substitute for $\hat{U}_j \hat{U}_{j+1}$.

B. Efficiency consideration

There are two questions that appear from the described closure procedure. First, how to choose sub-algebras so that the energy will be reduced efficiently? Second, is there a way to reduce the growth of the number of terms during generation of various closures?

To address the first question we can consider several heuristics for finding $\hat{U}$, they generally introduce the order dependence problem but with generating algebraic closures this dependency can be removed. The main ideas behind these heuristics are: 1) fermionic excitations (one-electron or mean-field picture), and 2) gradients along the $\hat{P}_k$'s (basis vectors in tangential space for the $SU(2^N)$ manifold).

For the second question, starting with a sub-algebra whose size is desirable to reduce, one can search for linear combinations of P_k's so that they commute with known symmetry operators. It is straightforward to show that if the Hamiltonian of the system has symmetry operators, $\{\hat{S}_k\}$ then a set of $\hat{P}_j$'s commuting with the symmetry operators form a sub-algebra, $\mathcal{S}_S$. It is enough to show that a commutator of any two elements of $\mathcal{S}_S$ also commutes with the symmetry operator and thus is in $\mathcal{S}_S$. To show this commutativity, we invoke the Jacobi condition for the commutator operation

$$[\hat{S}_k, [\hat{P}_i, \hat{P}_j]] + [\hat{P}_i, [\hat{P}_j, \hat{S}_k]] + [\hat{P}_j, [\hat{S}_k, \hat{P}_i]] = 0. \quad (7)$$

Since both $\hat{P}_i$ and $\hat{P}_j$ commute with $\hat{S}_k$ their commutator commutes with $\hat{S}_k$ as well. This consideration extends to any linear combination of $\hat{P}_k$'s since the commutator is a bilinear operation. Using the elements of $\mathcal{S}_S$ in the exponential map one can generate symmetry-adapted rotations that form a Lie sub-group. The number of generators in $\mathcal{S}_S$ cannot be larger than that in the original sub-algebra. Therefore, using symmetries like the number of electrons $\hat{N}_e$, the total electronic spin $\hat{S}^2$, and its z-projection $\hat{S}_z$, one can reduce the size of sub-algebras and construct sub-groups that will preserve the symmetries in unitary rotations.

Another approach to the Lie algebra reduction is through its formal decompositions (e.g. Levi decomposition).[17] Since we work with sub-algebras of the compact Lie algebra $su(2^N)$, one can use the following decomposition theorem: *any compact Lie algebra is a direct sum of the algebra center and simple sub-algebras.*[17] The algebra center commutes with all elements of the algebra, hence, based on Schur's lemma, action of the center elements on the wavefunction is equivalent to multiplication by a constant if the wavefunction is within a single irreducible representation. Since we exponentiate the algebra elements to obtain unitary transformations, one can remove the center elements from the construction if the wavefunction has all components within one irreducible representation of the center elements.

C. Two types of generators

A notion of a generator has some ambiguity when it is applied to Lie algebras. On the one hand, Lie algebras are linear spaces with $\hat{P}_k$'s as the basis elements (vectors). One can generate Lie groups via exponentiation of $\hat{P}_k$'s, and thus, it is reasonable to refer to all $\hat{P}_k$'s as generators of the associated Lie group. On the other hand, Lie algebras have a vector-vector multiplication operation, which is the commutator of two $\hat{P}_k$'s. Therefore, one can generate algebra basis vectors from a smaller subset of $\hat{P}_k$'s, which can be seen as the algebra generators. Both types of generators can be useful, therefore, to distinguish them we will refer to the latter generators as generators of the second type (generators II). To illustrate the difference between the two types of generators, let us consider a single-qubit case: there are three generators of the $SU(2)$ group, $\hat{x}$, $\hat{y}$, and $\hat{z}$, so that any single-qubit rotation can be written as

$$\hat{U}_1 = e^{i\theta_x \hat{x}} e^{i\theta_y \hat{y}} e^{i\theta_z \hat{z}} \qquad (8)$$

and two generators II, any two operators out of the triplet $(\hat{x}, \hat{y}, \hat{z})$ can generate the third operator via commutation. The generators II give rise to the Euler angle parametrization of single-qubit rotations

$$\hat{U}'_1 = e^{i\theta'_z \hat{z}} e^{i\theta_y \hat{y}} e^{i\theta_z \hat{z}}, \qquad (9)$$

note that repetition of some generators II is needed for a complete description of all rotations. This repetition can be considered as a compensation for a smaller set of exponentiated operators. It was shown previously, that the number of generators II scales only quadratically with N.[11] Essentially, to construct all $4^N - 1$ generators of $su(2^N)$ one needs only a subset of one- and two-qubit $\hat{P}_k$'s, for example, $\hat{x}$ and $\hat{z}$ operators for each qubit and pair of qubits. Similar considerations for the algebra of fermionic excitations shown that single and double excitations are sufficient to generate all possible excitations.[8] Note that due to necessity for repeated exponentiation of generators II, they create the most severe order problem in representation of unitaries, and we will not consider them further.

D. Unitary Coupled Cluster (UCC) Sub-algebras

Historically, the UCC method was formulated using an exponent of a sum for the unitary transformation,

$$\hat{U} = \exp\left[\sum_n \hat{T}_n - \hat{T}_n^\dagger\right]. \qquad (10)$$

Here, $\hat{T}_n$ are n-tuple excitations

$$\hat{T}_n = \sum_{ij...k, ab...c} t_{ij...k}^{ab...c} a_a^\dagger a_b^\dagger ... a_c^\dagger a_i a_j ... a_k, \qquad (11)$$

where we assume $i, j, ...k$ $(a, b, ...c)$ indices to correspond to occupied (unoccupied) molecular orbitals obtained in the Hartree-Fock method. The number of amplitudes $t_{ij...k}^{ab...c}$ grows combinatorially with the number of orbitals and electrons, therefore to have polynomial heuristics n is usually restricted to a fixed value, for example, $n = 1$ (singles, S)

$$\hat{T}_1 = \sum_{ia} t_i^a a_a^\dagger a_i. \qquad (12)$$

and 2 (doubles, D)

$$\hat{T}_2 = \sum_{ijab} t_{i,j}^{a,b} a_a^\dagger a_b^\dagger a_i a_j. \qquad (13)$$

This gives rise to the UCCSD parametrization,

$$\hat{U}_{SD} = \exp\left[\sum_{n=1}^2 \hat{T}_n - \hat{T}_n^\dagger\right]. \qquad (14)$$

All $\hat{T}_n$'s in Eq. (11) commute with each other componentwise, only addition of $\hat{T}_n^\dagger$ breaks down the commutativity. This poses a problem of presenting Eqs. (14) and (10) as a product of exponents of individual excitation/de-excitation pairs. The latter is needed for transferring the unitary transformation to a sequence of gates by compilers. Therefore, to implement UCCSD, one needs to do a Trotter approximation

$$\hat{U}_{SD} \approx \left[\prod_{ia} e^{t_i^a \hat{\kappa}_i^a / K} \prod_{ijab} e^{t_{ij}^{ab} \hat{\kappa}_{ij}^{ab} / K}\right]^K, \qquad (15)$$

where K is a finite number of Trotter steps, and

$$\hat{\kappa}_i^a = \hat{E}_i^a - \hat{E}_a^i = a_a^\dagger a_i - a_i^\dagger a_a \qquad (16)$$

$$\hat{\kappa}_{ji}^{ab} = \hat{E}_{ji}^{ab} - \hat{E}_{ab}^{ji} = a_a^\dagger a_b^\dagger a_i a_j - a_j^\dagger a_i^\dagger a_b a_a \qquad (17)$$

are anti-hermitized single and double excitations. Note that $\hat{\kappa}$ do not commute and order of terms in Eq. (15) matters for any finite K.

Using commutativity of $\hat{T}_n$ components one can "unitarize" the coupled cluster excitation operator

$$\hat{\Omega}_{SD} = \exp\left[\hat{T}_1 + \hat{T}_2\right] \qquad (18)$$

$$= \left[\prod_{ia} e^{t_i^a \hat{E}_i^a} \prod_{ijab} e^{t_{ij}^{ab} \hat{E}_{ij}^{ab}}\right]. \qquad (19)$$

as a product of exponents

$$\hat{U}_{dSD} = \left[\prod_{ia} e^{t_i^a \hat{\kappa}_i^a} \prod_{ijab} e^{t_{ij}^{ab} \hat{\kappa}_{ij}^{ab}}\right], \qquad (20)$$

which is called disentangled UCCSD (dUCCSD) in Ref. 8. dUCCSD will not require the Trotter approximation but its results will depend on the order of terms as was demonstrated numerically in Ref. 8.

To address the order problem in the UCC case we consider possible sub-algebras originating from sets of fermionic excitations/de-excitations $\hat{\kappa}$.

a. Individual unitary fermionic excitations/de-excitations: Here, we study sets of Pauli products obtained from transforming individual single and double fermionic excitations/de-excitations using the Jordan-Wigner (JW) transformation. This transformation was chosen for concreteness, similar results can be obtained using the parity or Bravyi-Kitaev transformations. It has been known that the individual fermionic excitation/de-excitation pairs produce sets of commuting Pauli products.[18] In case of singles and doubles these are

$$\hat{\kappa}_i^a = \frac{i}{2} \bigotimes_{k=i+1}^{a-1} \hat{z}_k (\hat{y}_i \hat{x}_a - \hat{x}_i \hat{y}_a) \tag{21}$$

$$\hat{\kappa}_{ij}^{ba} = \frac{i}{8} \bigotimes_{k=i+1}^{j-1} \hat{z}_k \bigotimes_{l=a+1}^{b-1} \hat{z}_l (\hat{x}_i \hat{x}_j \hat{y}_a \hat{x}_b + \hat{y}_i \hat{x}_j \hat{y}_a \hat{y}_b$$
$$+ \hat{x}_i \hat{y}_j \hat{y}_a \hat{y}_b + \hat{x}_i \hat{x}_j \hat{x}_a \hat{y}_b - \hat{y}_i \hat{x}_j \hat{x}_a \hat{x}_b - \hat{x}_i \hat{y}_j \hat{x}_a \hat{x}_b$$
$$- \hat{y}_i \hat{y}_j \hat{y}_a \hat{x}_b - \hat{y}_i \hat{y}_j \hat{x}_a \hat{y}_b). \tag{22}$$

One can continue this for higher excitations/de-excitations, where commutativity of Pauli terms will persist. The origin of this commutativity is in the even number of differences ($\hat{x}$ / $\hat{y}$ operators) in Pauli products surviving the subtraction of the de-excitation part from the excitation part (see more details in the Appendix of Ref. 18). All commuting Pauli products within an excitation/de-excitation set form an abelian sub-algebra, which obviously removes the order problem within this set.

All excitations/de-excitations commute with the number of electrons operator, $\hat{N}_e$, because they do not change the number of electrons (the number of a and $a^\dagger$ operators is the same). If one considers only $\hat{N}_e$-symmetry, the number of Pauli products within double excitations/de-excitations can be reduced to four terms which are mutually commutative

$$\hat{\xi}_{ij}^{ba} = \frac{i}{4} \bigotimes_{k=i+1}^{j-1} \hat{z}_k \bigotimes_{l=a+1}^{b-1} \hat{z}_l$$
$$(\hat{x}_i \hat{x}_j \hat{y}_a \hat{x}_b + \hat{y}_i \hat{x}_j \hat{y}_a \hat{y}_b - \hat{x}_i \hat{y}_j \hat{x}_a \hat{x}_b - \hat{y}_i \hat{y}_j \hat{x}_a \hat{y}_b), \tag{23}$$

$$\hat{\pi}_{ij}^{ba} = \frac{i}{4} \bigotimes_{k=i+1}^{j-1} \hat{z}_k \bigotimes_{l=a+1}^{b-1} \hat{z}_l$$
$$(\hat{x}_i \hat{y}_j \hat{y}_a \hat{y}_b + \hat{x}_i \hat{x}_j \hat{x}_a \hat{y}_b - \hat{y}_i \hat{x}_j \hat{x}_a \hat{x}_b - \hat{y}_i \hat{y}_j \hat{y}_a \hat{x}_b). \tag{24}$$

Single excitations are not reducible without violating commutativity with $\hat{N}_e$, higher excitations (n-tuple with $n > 1$) are reducible if only commutativity with $\hat{N}_e$ is required.

Operators that commute with the electron spin operators, $\hat{S}_z$ and $\hat{S}^2$, can be obtained by anti-hermitization of singlet spherical tensor operators. General spherical tensor operator $\hat{T}^{S,M}$ is defined as

$$[\hat{S}_\pm, \hat{T}^{S,M}] = \sqrt{S(S+1) - M(M \pm 1)} \hat{T}^{S,M\pm1}, \tag{25}$$

$$[\hat{S}_z, \hat{T}^{S,M}] = M \hat{T}^{S,M}, \tag{26}$$

where S and M are electron spin and its projection to the z-axis, respectively. Equation $\hat{S}^2 = \hat{S}_- \hat{S}_+ + \hat{S}_z(\hat{S}_z + 1)$ can be used to show that any singlet spherical tensor operator, $\hat{T}^{0,0}$ will commute with $\hat{S}_z$ and $\hat{S}^2$.

There are standard approaches for producing spherical tensor operators,[19] they involve very similar techniques to those used for generating spin-adapted configuration state functions.[20,21] Individual single excitations are not $\hat{T}^{0,0}$ operators, therefore, one needs to group more than one excitation to obtain the singlet operator

$$\hat{T}_{ia,\bar{i}\bar{a}}^{0,0} = \hat{\kappa}_i^a + \hat{\kappa}_{\bar{i}}^{\bar{a}}, \tag{27}$$

here and further $a(\bar{a})$ and $i(\bar{i})$ are spin-orbitals that have $\alpha(\beta)$ spin parts. For double and higher excitations/de-excitations, one can use seniority (the number of unpaired electrons created by the operator) to assess whether an individual excitation/de-excitation pair can be a singlet operator. Zero seniority always provides a positive answer, for example, zero-seniority double excitation/de-excitation gives

$$\hat{T}_{\bar{i}\bar{i}a\bar{a}}^{0,0} = \hat{\kappa}_{\bar{i}\bar{i}}^{a\bar{a}}. \tag{28}$$

For seniority two, one needs to combine two double excitations/de-excitations to obtain the singlet operator

$$\hat{T}_{\bar{i}\bar{i}a\bar{b},\bar{i}\bar{i}b\bar{a}}^{0,0} = \hat{\kappa}_{\bar{i}\bar{i}}^{a\bar{b}} + \hat{\kappa}_{\bar{i}\bar{i}}^{b\bar{a}}. \tag{29}$$

Generally, the seniority number is equal to the number of individual excitation/de-excitation pairs in construction of singlet operators. It is worth noting that in all odd excitations, due to nonzero seniority, individual fermionic excitations/de-excitations do not form proper spin-conserving operators. This illustrates that imposing symmetry constraints to some sub-algebras can give empty sets and remove these sub-algebras from consideration too prematurely.

b. All n-tuple excitations/de-excitations: Among of all sets with a fixed rank of excitation only singles form closed sub-algebras

$$[\hat{E}_r^s, \hat{E}_p^q] = \hat{E}_r^q \delta_{ps} - \hat{E}_p^s \delta_{rq}, \tag{30}$$

where we use $p, q, r, s, ...$ indices for both occupied and unoccupied orbitals. This equation can be generalized for $\hat{\kappa}_r^s$'s. Starting from double excitations, commutators produce higher excitations, and therefore, fixed-excitation algebras do not exist

$$[\hat{E}_{pq}^{rs}, \hat{E}_{tu}^{vw}] = \delta_{pw} \hat{E}_{tuq}^{rsv} - \delta_{pv} \hat{E}_{tuq}^{rsw} + ... \tag{31}$$

where $E_{tuq}^{rsv} = a_r^\dagger a_s^\dagger a_v^\dagger a_q a_u a_t$, and ... contain other triple excitations and all double and single excitations. This consideration gives the same conclusion for anti-hermitized excitation/de-excitation pairs, hence, results of any finite order disentangled UCC beyond singles (e.g. dUCCSDT) will be order dependent. Interestingly, even if one adds rotations within occupied and unoccupied subspaces, a finite operator rank is not enough to provide a closure, therefore, works using finite rank generalized UCC need Trotter steps as well.[6]

c. Orbital restricted subsets of excitations/de-excitations: One way to obtain closed algebras is to restrict the number of orbitals involved in possible excitations and de-excitations. If no restriction is made then operators involving all orbitals will be needed to close the algebra. Generally, one would not only need operators involving excitations from occupied to unoccupied orbitals and corresponding de-excitations but also rotations within occupied and unoccupied spaces.

E. Qubit Coupled Cluster (QCC) Sub-algebras

Another approach to construction of the unitary transformation is through a direct selection of Pauli products skipping the fermionic picture. Indeed, search for an element of the $SU(2^N)$ group minimizing the energy function [Eq. (1)] can be done using energy gradients with respect to individual amplitudes, τ_k, in the representation of Eq. (3)

$$\frac{\partial E}{\partial \tau_k}\Big|_{\boldsymbol{\tau}=0} = i \langle \bar{0}| [\hat{H}, \hat{P}_k] |\bar{0} \rangle, \tag{32}$$

In the QCC method[11] these gradients are evaluated using the qubit mean-field reference state, for simplicity we will take the Hartree-Fock Slater determinant as the reference state, $|\bar{0}\rangle$. In the qubit space, $|\bar{0}\rangle$ corresponds to a collinear product of qubits aligned with the z-axis. Such a reference state simplifies the selection of $\hat{P}_k$'s based on the energy gradient.[22]

a. Equivalence classes based on the energy gradients: We refer to the set of all $\hat{P}_k$'s with nonzero absolute energy gradients in Eq. (32) as the *direct interaction set* (DIS), denoted herein as $\mathcal{D}$. The DIS is formed by n_p nonoverlapping subsets (equivalence classes), $\mathcal{D} = \bigcup_{k=1}^{n_p} \mathcal{G}_k$, where each subset $\mathcal{G}_k$ contains $O(2^{N-1})$ $\hat{P}_k$'s of identical gradient magnitude.[22] The number of classes n_p for the qubit Hamiltonian scales linearly in the number of terms it contains in the fermionic representation, and hence efficient ranking of all $\hat{P}_k$'s in the DIS is accomplished by performing a gradient calculation only for a single representative $\hat{P}_k$ for each class, thereby ranking the complete DIS of cardinality $O(2^{N-1}n_p)$ with merely n_p gradient computations.

The nonzero gradient condition selects only $\hat{P}_k$'s that contain not higher than fermionic 2-body interactions due to 2-body terms in the Hamiltonian. $\hat{P}'_k$s from the same gradient class commute and thus form a sub-algebra. The sub-algebra of each nonzero gradient class is a linear subspace from which double fermionic excitation/de-excitation pairs can be built. Adding an extra condition of commutation with symmetry operators $\hat{N}_e$, $\hat{S}^2$, and $\hat{S}_z$ provides symmetry adapted combinations identical to those obtained in the double fermionic excitation/de-excitation case.

To go beyond fermionic doubles in this scheme one can generate new DIS by either modifying the Hamiltonian

or the reference state using unitary transformations built from high-gradient class $\hat{P}_k$'s. These approaches were used in recently suggested computational techniques, iterative QCC[22] and ADAPT-VQE.[12,13]

b. Sub-algebras based on anti-commuting sub-sets: To create sub-algebras that are very different from those obtained in the fermionic excitation/de-excitation case one needs to abandon symmetry considerations for a moment. Using DIS sets, it is possible to construct closed Lie sub-algebras starting with mutually anti-commutative Pauli products. Selecting P_k's from different gradient classes, one can form a mutually anti-commutative set $\{\hat{P}_i\}$, where $\hat{P}_i\hat{P}_j = -\hat{P}_j\hat{P}_i$. A finite sub-algebra originates from these anti-commuting terms if all product pairs $\hat{P}_i\hat{P}_j$ are added to the set. Due to the involutory property of $\hat{P}_k$'s, $\hat{P}_k^2 = \hat{1}$, all anti-commuting terms and their product pairs form a closed Lie sub-algebra

$$[\hat{P}_i, \hat{P}_j] = 2\hat{P}_i\hat{P}_j \tag{33}$$

$$[\hat{P}_i\hat{P}_j, \hat{P}_k] = 0, \tag{34}$$

$$[\hat{P}_i\hat{P}_j, \hat{P}_j] = 2\hat{P}_i, \tag{35}$$

$$[\hat{P}_i\hat{P}_j, \hat{P}_i] = -2\hat{P}_j, \tag{36}$$

$$[\hat{P}_i\hat{P}_j, \hat{P}_k\hat{P}_l] = 0, \tag{37}$$

$$[\hat{P}_i\hat{P}_j, \hat{P}_i\hat{P}_l] = 2\hat{P}_l\hat{P}_j, \tag{38}$$

$$[\hat{P}_i\hat{P}_j, \hat{P}_i\hat{P}_l] = 2\hat{P}_l\hat{P}_j, \tag{39}$$

$$[\hat{P}_i\hat{P}_j, \hat{P}_j\hat{P}_l] = 2\hat{P}_i\hat{P}_l, \tag{40}$$

$$[\hat{P}_i\hat{P}_j, \hat{P}_k\hat{P}_i] = 2\hat{P}_j\hat{P}_k, \tag{41}$$

$$[\hat{P}_i\hat{P}_j, \hat{P}_k\hat{P}_j] = 2\hat{P}_k\hat{P}_i. \tag{42}$$

The maximal size of these sub-algebras scales quadratically with the number of qubits because the maximal number of fully anti-commuting terms scales linearly with the number of qubits.[23] To keep the algebra compact the products $\hat{P}_i\hat{P}_j$ need to be included with the imaginary unit factor $i\hat{P}_i\hat{P}_j$, which changes the commutation relations non-essentially.

Interestingly, all sub-algebras constructed based on anti-commutative terms do not have finite symmetry adapted sub-algebras. Any attempt to find symmetrized linear combinations with respect to electron spin and number operators lead to empty sets. To obtain symmetrized operators one needs to combine multiple anti-commutativity-based sub-algebras into a composite algebra, this composite algebra will usually contain symmetry adapted sub-algebras. However, the symmetry adapted sub-algebras obtained this way are not different from those generated using symmetrization of fermionic excitations/de-excitations.

Another structural feature of the Pauli products algebra is that any sub-set of commuting operators naturally constitutes a Lie sub-algebra, and any fully anti-commuting subset can be organized into a Lie sub-algebra if all products are added, but if a subset contains commuting and anti-commuting elements then its

closure may require an exponential number of elements. A simple illustration of the last case is that a set of all one- and two-qubit operators for N-qubits can generate all $4^N - 1$ elements of the $su(2^N)$ algebra.

III. MODEL EXAMPLE

To illustrate two main steps in eliminating the order problem we consider the recently proposed model example of two electrons within a space of four spin-orbitals, two occupied $i, \bar{i}$ and two unoccupied $a, \bar{a}$.[8] Thus, an arbitrary state of two electrons can be written as

$$|\Psi\rangle = c_1|i\bar{i}\rangle + c_2|a\bar{i}\rangle + c_3|\bar{a}i\rangle + c_4|a\bar{a}\rangle \qquad (43)$$

where $|i\bar{i}\rangle, |a\bar{i}\rangle, |\bar{a}i\rangle, |a\bar{a}\rangle$ are four Slater determinants and c_i's are their normalized coefficients, $\sum_i |c_i|^2 = 1$. It is possible to construct three fermionic excitation/de-excitation operators: $\hat{\kappa}_{i\bar{i}}^{a\bar{a}}, \hat{\kappa}_{\bar{i}}^{\bar{a}}, \hat{\kappa}_i^a$. Interestingly, the order in which these three operators are placed in the disentangled UCC wavefunction can strongly affect the result. These differences are so dramatic that one particular order when applied to the physical vacuum state $|i\bar{i}\rangle$

$$|\Psi_{211}\rangle = e^{t_3 \hat{\kappa}_{i\bar{i}}^{a\bar{a}}} e^{t_2 \hat{\kappa}_{\bar{i}}^{\bar{a}}} e^{t_1 \hat{\kappa}_i^a} |i\bar{i}\rangle \qquad (44)$$

cannot represent an arbitrary state by varying the amplitudes t_i, while different orders like the following

$$|\Psi_{121}\rangle = e^{t_2 \hat{\kappa}_{\bar{i}}^{\bar{a}}} e^{t_3 \hat{\kappa}_{i\bar{i}}^{a\bar{a}}} e^{t_1 \hat{\kappa}_i^a} |i\bar{i}\rangle \qquad (45)$$

can. This order dependence is easy to understand if one creates the algebraic closure with respect to commutation for the set of the three fermionic operators. This closure results in the Lie algebra containing 8 operators:

$$\mathcal{A} = \{\hat{\kappa}_{i\bar{i}}^{a\bar{a}}, \hat{\kappa}_{\bar{i}}^{\bar{a}}, \hat{\kappa}_i^a, \hat{\kappa}_{i\bar{a}}^{a\bar{i}}, (\hat{n}_a - \hat{n}_i)\hat{\kappa}_{\bar{i}}^{\bar{a}}, (\hat{n}_{\bar{a}} - \hat{n}_{\bar{i}})\hat{\kappa}_i^a,$$
$$(\hat{n}_a - \hat{n}_i)^2 \hat{\kappa}_{\bar{i}}^{\bar{a}}, (\hat{n}_{\bar{a}} - \hat{n}_{\bar{i}})^2 \hat{\kappa}_i^a\}, \qquad (46)$$

where $\hat{n}_p = a_p^\dagger a_p$. This algebra can be decomposed as a direct sum of the center

$$\mathcal{C} = \{[1 - (\hat{n}_a - \hat{n}_i)^2]\hat{\kappa}_{\bar{i}}^{\bar{a}}, [1 - (\hat{n}_{\bar{a}} - \hat{n}_{\bar{i}})^2]\hat{\kappa}_i^a\} \quad (47)$$

and two simple algebras isomorphic to $su(2)$

$$\mathcal{A}_1 = \{(\hat{n}_{\bar{a}} - \hat{n}_{\bar{i}})\hat{\kappa}_i^a, (\hat{\kappa}_{i\bar{i}}^{a\bar{a}} + \hat{\kappa}_{i\bar{a}}^{a\bar{i}}), (\hat{n}_a - \hat{n}_i)^2 \hat{\kappa}_{\bar{i}}^{\bar{a}}\} \quad (48)$$

$$\mathcal{A}_2 = \{(\hat{n}_a - \hat{n}_i)\hat{\kappa}_{\bar{i}}^{\bar{a}}, (\hat{\kappa}_{i\bar{i}}^{a\bar{a}} - \hat{\kappa}_{i\bar{a}}^{a\bar{i}}), (\hat{n}_{\bar{a}} - \hat{n}_{\bar{i}})^2 \hat{\kappa}_i^a\}. \quad (49)$$

All elements of $\mathcal{C}$ zero any configuration with two electrons, hence, to avoid the order problem one needs to use only elements from $\mathcal{A}_1$ and $\mathcal{A}_2$.

Further reduction of the number of operators can be achieved by imposing the symmetry constraints to obtain symmetry adapted sub-algebras. All elements of $\mathcal{A}$ commute with $\hat{N}_e$ and $\hat{S}_z$, but to introduce the commutativity with $\hat{S}^2$ one needs to combine a few elements

of $\mathcal{A}$ together. The symmetry-adapted set of operators forming a Lie sub-algebra $\{\hat{A}_i\}$ is

$$\hat{A}_1 = \hat{\kappa}_i^a + \hat{\kappa}_{\bar{i}}^{\bar{a}} \qquad (50)$$

$$\hat{A}_2 = \hat{\kappa}_{i\bar{i}}^{a\bar{a}} \qquad (51)$$

$$\hat{A}_3 = (\hat{n}_a - \hat{n}_i)\hat{\kappa}_{\bar{i}}^{\bar{a}} + (\hat{n}_{\bar{a}} - \hat{n}_{\bar{i}})\hat{\kappa}_i^a \qquad (52)$$

$$\hat{A}_4 = (\hat{n}_a - \hat{n}_i)^2 \hat{\kappa}_{\bar{i}}^{\bar{a}} + (\hat{n}_{\bar{a}} - \hat{n}_{\bar{i}})^2 \hat{\kappa}_i^a. \qquad (53)$$

It is easy to show that all these operators are singlet spherical tensor operators. This algebra can be decomposed to the center

$$\hat{A}_C = [1 - (\hat{n}_a - \hat{n}_i)^2]\hat{\kappa}_{\bar{i}}^{\bar{a}} + [1 - (\hat{n}_{\bar{a}} - \hat{n}_{\bar{i}})^2]\hat{\kappa}_i^a \quad (54)$$

and the simple algebra isomorphic to $su(2)$

$$\mathcal{A}_S = \{\hat{A}_2, \hat{A}_3/2, \hat{A}_4/2\}. \qquad (55)$$

$\hat{A}_C$ zeroes any configuration in Eq. (43), and thus one can use only elements of $\mathcal{A}_S$ to build the singlet wavefunction without the order dependence as

$$\hat{U}(\boldsymbol{\tau}) |i\bar{i}\rangle = \prod_{j=2}^4 e^{\tau_j \hat{A}_j} |i\bar{i}\rangle, \qquad (56)$$

where $\boldsymbol{\tau} = \{\tau_j\}$ are three optimization parameters (note that their optimal values can depend on the order of the exponents in the product). As a numerical test for this ansatz we optimized $\hat{U}(\boldsymbol{\tau})$ to reproduce the wavefunction

$$|\tilde{\Psi}\rangle = \frac{1}{\sqrt{2}}\left(|a\bar{i}\rangle + |i\bar{a}\rangle\right), \qquad (57)$$

which was shown to be unreachable using Eq. (44).[8] In addition, we tested all 6=3! different orderings of exponents in Eq. (56) (e.g. $e^{\tau_4 \hat{A}_4} e^{\tau_3 \hat{A}_3} e^{\tau_2 \hat{A}_2}$) and obtained $|\langle \tilde{\Psi} | \hat{U}(\boldsymbol{\tau}) |i\bar{i}\rangle| = 1$ for all of them.

IV. CONCLUSIONS

The order problem appears when one uses a partial product of non-commuting elements of the $SU(2^N)$ Lie group to construct a unitary transformation in the VQE approach. Here, we have shown that elimination of the order dependence can be done by using Lie sub-groups generated from corresponding Lie sub-algebras. Sub-algebras are obtained by considering all commutators of some initial sub-set generated using heuristics for energy minimization. Sub-algebra construction increases the number of elements in the unitary transformation, to reduce the number of parameters for optimization without breaking the closure, we proposed reduction based on the system symmetry and analysis of the sub-algebra center. This reduction builds symmetry-adapted linear combinations of sub-algebra elements that commute with the symmetry operators.

Some caution needs to be exercised in application of the proposed procedures. To eliminate the order problem completely, one needs to generate a sub-algebra containing all generators important for the energy lowering. This sub-algebra can be quite large so that its symmetry-adapted reduction is still large. The symmetry adaptation, even though reduces the number of parameters to optimize, constructs group elements that can require deep circuits to be implemented. This aspect requires further careful investigation. On the other hand, if one does not create large sub-algebras, imposing symmetrization can lead to empty sets of symmetrized generators (e.g., single fermionic excitation/de-excitation sets) and as a result removal of important operators.

V. ACKNOWLEDGEMENTS

[*] artur.izmaylov@utoronto.ca

[1] A. Peruzzo, J. McClean, P. Shadbolt, M.-H. Yung, X.-Q. Zhou, P. J. Love, A. Aspuru-Guzik, and J. L. O'Brien, Nat. Commun. **5**, 4213 (2014).

[2] R. Gilmore, in *Lie Groups, Physics, and Geometry: An Introduction for Physicists, Engineers and Chemists* (Cambridge University Press, Cambridge, UK, 2008) p. 57.

[3] A. W. Knapp, in *Lie groups beyond an introduction* (Second Edition, Progress in Mathematics, Vol. 140, Birkhuser, Boston, 2002) p. 259.

[4] R. Gilmore, in *Lie Groups, Physics, and Geometry: An Introduction for Physicists, Engineers and Chemists* (Cambridge University Press, Cambridge, UK, 2008) p. 108.

[5] Note that quantum computers cannot perform $\exp\left(i\tau_k \hat{P}_k\right)$ as elementary circuit operations (gates) but there are compilers that present such operations as sequences of universal gates containing not more than polynomial in N number of gates. Thus, we assume that polynomial in the number of P_k's algorithms is a desirable goal.

[6] J. Lee, W. J. Huggins, M. Head-Gordon, and K. B. Whaley, J. Chem. Theory Comput. **15**, 311 (2019).

[7] I. Sokolov, P. K. Barkoutsos, P. J. Ollitrault, D. Greenberg, J. Rice, M. Pistoia, and I. Tavernelli, arXiv.org (2019), arXiv:1911.10864 [quant-ph].

[8] F. A. Evangelista, G. K.-L. Chan, and G. E. Scuseria, J. Chem. Phys. **151**, 244112 (2019).

[9] W. Mizukami, K. Mitarai, Y. O. Nakagawa, T. Yamamoto, T. Yan, and Y.-y. Ohnishi, arXiv e-prints (2019), arXiv:1910.11526 [cond-mat.str-el].

[10] C. Hempel, C. Maier, J. Romero, J. McClean, T. Monz, H. Shen, P. Jurcevic, B. P. Lanyon, P. Love, R. Babbush, A. Aspuru-Guzik, R. Blatt, and C. F. Roos, Phys. Rev. X **8**, 31022 (2018).

[11] I. G. Ryabinkin, T.-C. Yen, S. N. Genin, and A. F. Izmaylov, J. Chem. Theory Comput. **14**, 6317 (2018).

[12] H. R. Grimsley, S. E. Economou, E. Barnes, and N. J. Mayhall, Nat. Commun. **10**, 3007 (2019).

[13] H. L. Tang, E. Barnes, H. R. Grimsley, N. J. Mayhall, and S. E. Economou, arXiv.org (2019), 1911.10205v1.

[14] H. R. Grimsley, D. Claudino, S. E. Economou, E. Barnes, and N. J. Mayhall, arXiv.org (2019), 1910.10329v1.

[15] A. Tranter, P. J. Love, F. Mintert, N. Wiebe, and P. V. Coveney, Entropy **21**, 1218 (2019).

[16] R. Gilmore, *Lie Groups, Physics, and Geometry: An Introduction for Physicists, Engineers and Chemists* (Cambridge University Press, 2008).

[17] A. O. Barut and R. Raczka, in *Theory of Group Representations and Applications* (Polish Scientific Publisher, Warszawa, Poland, 1980) p. 17.

[18] J. Romero, R. Babbush, J. R. McClean, C. Hempel, P. J. Love, and A. Aspuru-Guzik, Quantum Sci. and Technol. **4**, 014008 (2018).

[19] T. Helgaker, P. Jørgensen, and J. Olsen, *Molecular Electronic-Structure Theory* (John Wiley & Sons, Ltd, Chichester, UK, 2000).

[20] X. Li and J. Paldus, Theor. Chem. Acc. **133**, 1467 (2014).

[21] I. Shavitt, Int. J. Quant. Chem. **14**, 5 (2009).

[22] I. G. Ryabinkin, R. A. Lang, S. N. Genin, and A. F. Izmaylov, J. Chem. Theory Comput. **16**, 1055 (2020).

[23] A. Zhao, A. Tranter, W. M. Kirby, S. F. Ung, A. Miyake, and P. Love, arXiv e-prints (2019), arXiv:1908.08067 [quant-ph].