Quantum and quantum-inspired algorithms for the electronic structure problem

by

Robert A. Lang

A thesis submitted in conformity with the requirements
for the degree of Doctor of Philosophy

Department of Chemistry
University of Toronto

Quantum and quantum-inspired algorithms for the electronic structure problem

Robert A. Lang
Doctor of Philosophy
Department of Chemistry
University of Toronto
2024

Abstract

Ab initio quantum chemistry simulations are crucially useful for the accurate prediction and understanding of chemical phenomena. The quality of such simulations rely heavily on precise solutions to the electronic structure problem. To facilitate the approximate yet efficient solving of the electronic structure problem, many methods rely on a set of problem assumptions, such as the validity of truncations of the fermionic excitation hierarchy, or active space selections. In this work, we describe black-box algorithms for the electronic structure problem, namely the iterative qubit coupled cluster (iQCC) algorithm along with a few of its variants and extensions. The iQCC method is a variational adaptive technique, where an iteratively updated effective Hamiltonian directly guides the selection of transformation generators of the next iteration. We describe the algorithm in detail, and introduce variations of the algorithm to alleviate the proliferation of terms in the effective Hamiltonians. Finally, we describe how the algorithm can be modified to perform simultaneuous determination of ground and excited states. These methods are amenable to implementation using near-term quantum computing devices, and with certain restrictions on meta-parameters, can be implemented as fully classical "quantum-inspired" algorithms.

List of Acronyms

AO atomic orbital

BCH Baker-Campbell-Hausdorff

BFGS Broyden-Fletcher-Goldfarb-Shanno

BK Bravyi-Kitaev

CAS complete active space

CC coupled cluster

CCSD coupled cluster with single and doubles

CI configuration interaction

DIS direct interaction space

FCI full configuration interaction

GM-iQCC growth mitigated iterative qubit coupled cluster

HF Hartree-Fock

ILC involutory linear combination

iQCC iterative qubit coupled cluster

iQCC-ILC iterative qubit coupled cluster with involutory linear combinations

JW Jordan-Wigner

MO molecular orbital

MRSA-iQCC multireference state-averaged iQCC

PES potential energy surface

QMF qubit mean-field

QPE quantum phase estimation

RHF restricted Hartree Fock

UCC unitary coupled cluster

UCCSD unitary coupled cluster with singles and doubles

VQE variational quantum eigensolver

Acknowledgements

Firstly, a big thank you to NSERC for graduate student funding.

This thesis would not have been possible without the help of many amazing people, and it brings me great pleasure to thank a few in particular. I am especially grateful to my thesis supervisor, Prof. Artur Izmaylov, for his support, patience, and teaching me more than I could possibly capture here. I firmly believe his mentorship has provided me with the necessary foundations for an effective scientific career.

Our group has seen quite a few changes from when I was first welcomed in, and I am grateful to have had the privilege to interact and work with so many outstanding former and current students and postdocs. In particular, Dr. Ilya Ryabinkin and Tzu-Ching (Thomson) Yen, who were incredibly helpful when I first joined and a joy to collaborate with. To name a few others: Dr. Seonghoon Choi, Dr. Ignacio Loaiza, Smik Patel, and Praveen Jayakumar for fascinating discussions, all of the other exceptional people I had the pleasure of befriending in the Chemical Physics Theory Group, and all of the teachers who inspired me. I am thankful to my committee members, Prof. Paul Brumer and Prof. Alán Aspuru-Guzik, for their valuable insights during our meetings. I am also grateful to Dr. Junan Lin for his valuable feedback while proof-reading this work. Of course, any shortcomings of this work are solely the fault of the author.

I also owe gratitude to all of the talented people on the Amazon Braket team at AWS Quantum Technologies I had the opportunity to work with during my time there. In particular, Cody Wang, Dr. Cedric Yen-Yu Lin, Dr. Eric Kessler, along with Xanadu members Dr. Josh Izaac, Dr. Nathan Killoran, and Dr. Soran Jahangiri, for their patience, guidance, and frequent valuable input.

This list would be regrettably incomplete without mention of Prof. Tao (Toby) Zeng. It's difficult to say where I might be had we not crossed paths back at Carleton, but I am confident that none of this would have come to fruition. I couldn't have asked for a better scientific mentor so early on in my journey.

Lastly, but certainly not least, I am forever indebted to my parents. Their unwavering support has granted me the freedom to pursue my endeavours wholeheartedly. Love you, Mom and Dad.

I have learned so much from you all!

Contents

Preface

This thesis is based on selected works produced as a result of my research conducted at the University of Toronto. Many of the results presented within are available in previously published, or soon to be published papers:

[1]: Ilya G. Ryabinkin, Robert A. Lang, Scott N. Genin, and Artur F. Izmaylov. Iterative qubit coupled cluster approach with efficient screening of generators. *J. Chem. Theory Comput.*, 16(2):1055–1063, 2020.

[2]: Robert A. Lang, Ilya G. Ryabinkin, and Artur F. Izmaylov. Unitary transformation of the electronic Hamiltonian with an exact quadratic truncation of the Baker-Campbell-Hausdorff expansion. *J. Chem. Theory Comput.*, 17(1):66–78, 2021.

[3]: Robert A. Lang, Aadithya Ganeshram, and Artur F. Izmaylov. Growth reduction of similarity-transformed electronic Hamiltonians in qubit space. *J. Chem. Theory Comput.*, 19(19):6656–6667, 2023.

[4]: Robert A. Lang and Artur F. Izmaylov. A multireference state-averaged iterative qubit coupled cluster method for ground and excited state electronic energies. *Under preparation*

In the first chapter of this thesis, an introduction is provided, including the necessary formal background to the electronic structure problem, and aspects to addressing it in the setting of qubit algebra. Chapter 2 corresponds to Reference 1, which introduces the iterative qubit coupled cluster (iQCC) algorithm for ground state electronic structure calculations. I.G.R conceptualized the main methodology and produced the numerical results, the present author formulated the generator screening procedure, and all authors collaborated on the theoretical background and methodology. Chapters 3 and 4 review methods to improve upon the resource efficiency of the iQCC algorithm, and are largely summaries of References 2 and 3, respectively. I.G.R provided theoretical insights and aided in divising the algorithm for producing anticommuting sets used in Reference 2. A.G. performed exploratory numerical simulations during the devising of the method put forth in Reference 3.

In Chapter 5, the iQCC formalism is generalized in order to tackle simultaneous electronic ground and excited state estimations, and corresponds to Reference 4. The first author of References 2, 3 and 4 conducted the algorithm implementations, produced the published numerical results, and performed the mathematical derivations; all authors collaborated on the theoretical background.

Chapter 1

Introduction

The accurate simulation of polyatomic systems such as those including molecules and solids, through first principle methods, has proven to be indispensable across various scientific and industrial domains. These quantum chemical simulations can offer invaluable insights and can direct the design of novel molecules, and the optimization of chemical processes. Such applications range from drug design, solar light harvesting, industrial catalysts, agriculturally relevant molecules, and many more. Beyond the industrial significance of predictively accurate simulations, they can also contribute significantly to our scientific understanding of many complex chemical processes, such as those occurring in biological systems.

Commensurate with the significance of accurate first-principle simulations of polyatomic systems is their general computational hardness. At the heart of quantum chemistry is the electronic structure problem, involving the accurate solving of eigenstates and eigenenergies of the electronic Hamiltonian. The field of electronic structure has matured for over a century, and it continues to extend its scope via theoretical breakthroughs and advancements in computational capabilities. The ability to simulate systems of increasing complexity has been a hallmark of progress in this field. Despite several breakthroughs, a universally efficient blackbox prescription for solving the electronic structure problem across diverse instances remains elusive. As such, there exist many molecular systems of industrial or scientific relevance which are out of scope of state-of-the-art classical electronic structure methods. It was over forty years ago when the idea of simulating quantum mechanical systems is most naturally accomplished using a computer which itself behaves quantum mechanically. In more recent years, we have seen the inception of such quantum computing devices, and their rapid improvement in computational capabilities. However, the bleeding edge in quantum computing devices are still far from idealized, having various limitations, hindering their application to classically intractable instances of the electronic structure problem.

This work largely consists of a description of developments made to the iterative qubit coupled cluster (iQCC) algorithm, a robust electronic structure method for ground state estimation.[1] In light of the limited resources of existing and near-term quantum hardware,

the iQCC method was initially proposed as a near-term quantum algorithm, with the goal of bringing the solving of challenging instances of the electronic structure problem closer to realization utilizing quantum computation. Due to flexibilities in the implementation of the iQCC algorithm, it has also seen success as a "quantum-inspired" fully classical algorithm.[5] In this work, we review the developments of the iQCC algorithm, along with it's variants, in a manner which is implementation-agnostic. The numerical assessments included in this work are discussed from the standpoint of implementation as both a near-term quantum algorithm and a fully classical algorithm, with their differences discussed when appropriate.

The remainder of this chapter is structured as follows. In Section 1.1, we formally introduce the electronic structure problem, and review its formulation in the second quantization. In Section 1.2, we provide a very brief review of the necessary aspects of quantum computing used in the remainder of this work. In Section 1.3, we review the concept of fermion-to-qubit mappings, used to formulate the electronic structure problem in the algebra of quantum computers. In Section 1.4, we review one of the most widely explored frameworks for addressing the electronic structure problem leveraging near-term quantum computers; the variational quantum eigensolver (VQE). Finally, in Section 1.5, we review some of the basic properties of an "adaptive" quantum circuit ansatz, which leads naturally into the introduction of the iQCC algorithm presented in Chapter 2.

1.1 Electronic structure

Understanding the electronic structure of atoms, molecules, and solids lies at the heart of quantum chemistry, serving as a cornerstone for studying their properties and behavior. The electronic structure provides insights into a myriad of chemical phenomena, ranging from molecular bonding to spectroscopic transitions. The insights derived from electronic structure calculations have far-reaching applications across diverse fields, ranging from pharmaceutical drug discovery and novel material design to agriculture and environmental science. By harnessing the predictive power of quantum chemistry, researchers and practitioners can address critical challenges and innovate in various domains. However, despite their versatility and utility, electronic structure calculations face significant challenges when applied to large and intricately correlated systems. As the size and complexity of molecular and material systems increase, the computational demands of accurately modeling their electronic structure grow exponentially. The computational cost associated with solving the time-independent electronic Schrödinger equation to sufficient accuracy often exceeds the capabilities of conventional computing resources, limiting the feasibility of accurate predictions for large-scale, strongly correlated systems. Within this section, we review the formal setting of the electronic structure problem. In Section 1.1.1 we introduce the time-independent electronic Schrödinger equation, the central equation in the electronic structure problem. In Section 1.1.2, we move to the second quantized formulation of the electronic

structure problem.

1.1.1 The electronic Schrödinger equation

The quantum Hamiltonian for a molecular system consisting of N_e electrons with spatial coordinates $\{\boldsymbol{r}_i\}_{i=1}^{N_e}$ and N_n nuclei with coordinates $\{\boldsymbol{R}_I\}_{I=1}^{N_n}$, masses $\{M_I\}_{I=1}^{N_n}$, and atomic numbers $\{Z_I\}_{I=1}^{N_n}$, is expressed as

$$\hat{H} = -\frac{\hbar^2}{2m_e}\sum_{i}^{N_e}\nabla_i^2 - \frac{\hbar^2}{2}\sum_{I}^{N_n}\frac{1}{M_I}\nabla_I^2 + \frac{e^2}{4\pi\epsilon_0}\left(\frac{1}{2}\sum_{i\neq j}^{N_e}\frac{1}{|\boldsymbol{r}_i-\boldsymbol{r}_j|} + \frac{1}{2}\sum_{I\neq J}^{N_n}\frac{Z_I Z_J}{|\boldsymbol{R}_I-\boldsymbol{R}_J|} - \sum_{i}^{N_e}\sum_{I}^{N_n}\frac{Z_I}{|\boldsymbol{r}_i-\boldsymbol{R}_I|}\right), \tag{1.1}$$

where ∇_p^2 is the Laplace operator taken with respect to the spatial coordinates of particle p. The first two terms denote the contributions from the electronic kinetic energy $\hat{T}_e$ and nuclear kinetic energy $\hat{T}_n$, respectively. The following terms arise from the potential energy arising from the interactions of particles, which are the electron-electron repulsion $\hat{V}_{ee}$, the nuclear-nuclear repulsion $\hat{V}_{nn}$, and the electron-nuclear attraction $\hat{V}_{en}$ respectively. Hence the molecular Hamiltonian may be concisely written as

$$\hat{H} = \hat{T}_e + \hat{T}_n + \hat{V}_{ee} + \hat{V}_{nn} + \hat{V}_{en}. \tag{1.2}$$

For the remainder of the work presented, the Hamiltonian is considered in atomic units where the reduced Planck constant $\hbar$, the electron mass m_e, the elemental charge e, and $4\pi\epsilon_0$ (where ϵ_0 is the vacuum permittivity), are all set to one.

To solve for strictly the electronic degrees of freedom at a fixed nuclear geometry, one considers the electronic Hamiltonian depending only *parameterically* on the nuclear coordinates $\boldsymbol{R} \equiv \{\boldsymbol{R}_I\}_{I=1}^{N_n}$,

$$\hat{H}_e(\boldsymbol{r};\boldsymbol{R}) = \hat{T}_e(\boldsymbol{r}) + \hat{V}_{ee}(\boldsymbol{r}) + \hat{V}_{en}(\boldsymbol{r};\boldsymbol{R}) + V_{nn}(\boldsymbol{R})\hat{I}, \tag{1.3}$$

where $\hat{I}$ is the identity operator, and $V_{nn}(\boldsymbol{R})$ is a trivially evaluated constant. The electronic structure problem amounts to finding the solutions to the time-independent Schrödinger equation for Eq.(1.3) at a given arrangement of nuclei $\boldsymbol{R}$,

$$\hat{H}_e(\boldsymbol{R})\,|\Psi_i[\boldsymbol{R}]\rangle = E_i(\boldsymbol{R})\,|\Psi_i[\boldsymbol{R}]\rangle\,. \tag{1.4}$$

By solving for the electronic eigenenergy E_i on a set of grid points in nuclear configuration space $\{\boldsymbol{R}\}$, one obtains a potential energy surface (PES) of electronic state $|\Psi_i\rangle$. Formally, exact solution to Eq.(1.4) requires the search over the space of normalized functions of $\boldsymbol{r}$, requiring an infinite $3N_e$-dimensional grid of infinitismally spaced points to represent the

electronic wavefunction in real or momentum space. Hence, the first step to computationally solving the electronic structure problem is a discretization of the electronic space to a finite basis. For molecular systems, it is often reasonable to utilize a set of atom-centered functions resembling the one-electron atomic stationary solutions. Such atomic orbitals (AOs) are often approximated using gaussian functions for more efficient integral calculation.[6] The basis set utilized in the polyatomic problem are then taken to be linear combinations of AOs, with the optimal set of linear combinations obtained via e.g., the Hartree Fock procedure.[6] As it is comprised of fermions, the electronic wavefunction $\Psi(\boldsymbol{r})$ must be totally antisymmetric with respect to the transpositions of electronic coordinates, that is

$$\Psi(\boldsymbol{r}_1, \ldots, \boldsymbol{r}_i, \ldots, \boldsymbol{r}_j \ldots, \boldsymbol{r}_{N_e}) = -\Psi(\boldsymbol{r}_1, \ldots, \boldsymbol{r}_j, \ldots, \boldsymbol{r}_i \ldots, \boldsymbol{r}_{N_e}) \tag{1.5}$$

for arbitrary pair $\boldsymbol{r}_i \neq \boldsymbol{r}_j$. Considering a basis of $\{\chi_p\}_{p=1}^{M}$ one-electron orbitals, one can construct an arbitrary antisymmetric N_e-electron function by considering the basis of N_e-electron Slater determinants formable in $\{\chi_p\}_{p=1}^{M}$. We denote a Slater determinant with $N_e \leq M$ electrons in orbitals $\{\chi_{n_1}, \chi_{n_2}, \ldots, \chi_{n_{N_e}}\}$ as $\phi_{n_1,\ldots,n_{N_e}}(\boldsymbol{r})$, which is explicitly expressed as

$$\phi_{n_1,\ldots,n_{N_e}}(\boldsymbol{r}) = \frac{1}{\sqrt{N_e!}} \det \begin{bmatrix} \chi_{n_1}(\boldsymbol{r}_1) & \chi_{n_1}(\boldsymbol{r}_2) & \cdots & \chi_{n_1}(\boldsymbol{r}_{N_e}) \\ \chi_{n_2}(\boldsymbol{r}_1) & \chi_{n_2}(\boldsymbol{r}_2) & \cdots & \chi_{n_2}(\boldsymbol{r}_{N_e}) \\ \vdots & \vdots & \ddots & \vdots \\ \chi_{n_{N_e}}(\boldsymbol{r}_1) & \chi_{n_{N_e}}(\boldsymbol{r}_2) & \cdots & \chi_{n_{N_e}}(\boldsymbol{r}_{N_e}) \end{bmatrix}, \tag{1.6}$$

where $\det A$ is denotes the determinant of matrix A. Consideration of all $\binom{M}{N_e}$ Slater determinants yields a complete N_e-electron basis for $|\Psi(\boldsymbol{r}\rangle)$ as antisymmetrized N_e-times products of one-electron basis $\{\chi_p\}_{p=1}^{M}$. Within the next section, we describe an alternative representation of the electronic structure problem, which as we will see, becomes quite convenient for our purposes of employing quantum computers towards this task.

1.1.2 Fermionic algebra

Expressing the electronic wavefunction by the explicit use of Slater determinants is referred to as the first quantization picture. The manipulation of Slater determinants explicitly in solving the many-electron problem is often overly tedious (however there exist use-cases where first quantization possesses advantages[7, 8]), and many electronic structure methods avoid formulation in first quantization. More commonly, including in this work, the electronic wavefunction is expressed in a different yet equivalent representation referred to as second quantization. In second quantization, a Slater determinant is encoded as a *Fock*

occupation vector,

$$|n_1, \ldots, n_M\rangle\,, \quad n_p \in \{0, 1\}, \tag{1.7}$$

where $n_p = 1$ denotes the occupation of orbital χ_p by an electron, and $n_p = 0$ specifies the orbital is empty (virtual). Any fermionic occupation vector of N_e electrons expressed in the spin-orbital basis $\{\chi_p\}_{p=1}^{M}$ can be expressed in terms of a product string of fermionic creation operators, acting on the physical electronic vacuum $|\text{vac}\rangle$,

$$|n_1, \ldots, n_M\rangle = \prod_{\substack{p=1 \\ n_p=1}}^{N_e} \hat{a}_p^\dagger \,|\text{vac}\rangle\,, \tag{1.8}$$

where $\hat{a}_p^\dagger$ creates an electron in the p^{th} spin orbital. The vacuum is defined to vanish under the action of any fermionic annihilation operator, i.e., $\hat{a}_p\,|\text{vac}\rangle = 0$ for all p. The antisymmetry of the electronic state is implicitly enforced by the anticommutation rules of the fermionic operators,

$$\{\hat{a}_p^\dagger, \hat{a}_q^\dagger\} = \{\hat{a}_p, \hat{a}_q\} = 0 \tag{1.9}$$

$$\{\hat{a}_p^\dagger, \hat{a}_q\} = \delta_{pq}, \tag{1.10}$$

where the anticommutator is defined as $\{\hat{A}, \hat{B}\} = \hat{A}\hat{B} + \hat{B}\hat{A}$. Any operators acting on the electronic space are likewise resolved in terms of the fermionic creation and annihilation operators, with the electronic structure Hamiltonian being written as

$$\hat{H} = \sum_{pq} h_{pq}\hat{a}_p^\dagger\hat{a}_q + \frac{1}{2}\sum_{pqrs} g_{pqrs}\hat{a}_p^\dagger\hat{a}_q^\dagger\hat{a}_s\hat{a}_r + V_{nn}\hat{I} \tag{1.11}$$

where V_{nn} is the constant arising from the nuclear repulsion, and h_{pq} and g_{pqrs} are one- and two-electron integrals, respectively,

$$h_{pq} = \int d\boldsymbol{r_1}\chi_p^*(\boldsymbol{r_1})\left(-\frac{1}{2}\nabla^2 - \sum_{I=1}^{N_n}\frac{Z_I}{|\boldsymbol{r}_1 - \boldsymbol{R}_I|}\right)\chi_q(\boldsymbol{r_1}) \tag{1.12}$$

$$g_{pqrs} = \int d\boldsymbol{r_1}d\boldsymbol{r_2}\frac{\chi_p^*(\boldsymbol{r}_1)\chi_q^*(\boldsymbol{r}_2)\chi_r(\boldsymbol{r}_1)\chi_s(\boldsymbol{r}_2)}{|\boldsymbol{r}_1 - \boldsymbol{r}_2|}. \tag{1.13}$$

An arbitrary N_e-electron state defined over M spin-orbitals may be written as

$$|\Psi\rangle = \sum_{i=1}^{\binom{M}{N_e}} C_i \left|n_1^{(i)}, \ldots, n_M^{(i)}\right\rangle, \quad \sum_{p=1}^{M} n_p^{(i)} = N_e. \tag{1.14}$$

The exact ground state in the finite orbital basis may be obtained by free optimization of coefficients C_i, and such state is referred to as the full configuration interaction (FCI) state. Due to the exponentially scaling number of configurations in Eq.(1.14), such a procedure is only feasible for very modestly sized basis sets. The goal of any approximate electronic structure method is to parameterize the trial wavefunction sufficiently close to the FCI solution while using an efficient amount of compute resources. To this end, there have been many decades of rich development to efficiently and accurately solve the electronic structure problem. Indeed, coupled with the improvement of available classical computing resources, mean-field procedures such as Hartree-Fock (HF) and density functional theory are routinely applied to systems in basis sets of up to $\sim 10,000$ spin-orbitals, and correlated methods with higher polynomial scalings such as coupled cluster (CC) and configuration interaction (CI) can be used to address systems with thousands of spin-orbitals using standard codes.[9] However, there exist many chemical settings where the efficient classical methodologies fail to achieve satisfactory accuracy. Examples of such difficult scenarios include molecular systems possessing multiple stretched bonds,[10] and systems containing highly degenerate orbital functions, such as those found in transition metal complexes.[11] To address systems currently outside the scope of accurate description by state-of-the-art classical methods, the advent of commercially available quantum computers presents itself as a promising path for new electronic structure methods development.[12, 13] In the next section, we review the necessary mathematical background of gate-based quantum computing and necessary steps towards their application to the electronic structure problem.

1.2 Quantum computing

Mathematically, a quantum computer consists of a collection of two-level systems which can be manipulated systematically. A qubit is defined as a two-level quantum system, living in complex Hilbert space $\mathbb{C}^2$, and it's pure states can be expressed generally as $\alpha |0\rangle + \beta |1\rangle$ with $|\alpha|^2 + |\beta|^2 = 1$. The single qubit basis states $\{|0\rangle, |1\rangle\}$ are often referred to as the "measurement" basis or "computational basis", as it is the basis which is natively measured in by quantum devices. Physically, such states may correspond to the direction of intrinsic particle spin measured along the z axis, the observation of a photon in a bosonic mode, etc., depending on the quantum hardware implementation. A quantum computer with a register of N_q qubits then lives in the tensor product space $\mathbb{C}^2 \otimes \ldots \otimes \mathbb{C}^2 = \mathbb{C}^{2^{N_q}}$, with general state written in the N_q-qubit computational basis

$$|\Psi_{\text{qc}}\rangle = \sum_{\boldsymbol{z} \in \{0,1\}^{\otimes N_q}} C_{\boldsymbol{z}} \left| z_1, \ldots, z_{N_q} \right\rangle \tag{1.15}$$

where $\boldsymbol{z} = (z_1, \ldots, z_{N_q})$ with $z_p \in \{0, 1\}$, and shorthand notation $\left| z_1, \ldots, z_{N_q} \right\rangle = |z_1\rangle \otimes \ldots \otimes \left| z_{N_q} \right\rangle$ is used. From Eq.(1.15), a very clear advantage of utilizing a quantum register to

represent a quantum state of interest is the memory required over it's classical counterpart. A quantum computer naturally encodes a set of 2^{N_q} coefficients $C_{\boldsymbol{z}}$ with N_q qubits, as opposed to requiring $O(2^{N_q})$ classical memory. In order to utilize the quantum computer to accomplish a computational task, one must have knowledge on how to prepare the target state starting from a trivial starting point, e.g.

$$|\Psi_{\text{qc}}\rangle = \hat{U}\,|0,\ldots,0\rangle\,, \tag{1.16}$$

where $\hat{U}$ is a unitary operation, $\hat{U}^\dagger\hat{U} = \hat{1}$, implementable on the quantum device. A given quantum device has a set of universal one- and two-qubit unitary operations, referred to as universal gate set, from which $\hat{U}$ may be decomposed into and performed on the quantum register. The set of N_q qubit unitary operations forms Lie group $U(2^{N_q}) \cong U(1)\times SU(2^{N_q})$. Since the $U(1)$ freedom amounts to an immeasurable phase, we refer to $SU(2^{N_q})$ as the group of interest when discussing operations implementable on the quantum register. When reasoning about unitary operations, it is often natural to consider their anti-Hermitian generators through the exponential map, $\hat{U} = e^{\hat{G}}$, where $\hat{G} = -\hat{G}^\dagger$, i.e., the $\mathfrak{su}(2^{N_q})$ algebra. The single qubit $\mathfrak{su}(2)$ algebra has the $\{-i\hat{x}, -i\hat{z}, -i\hat{y}\}$ Pauli matrices as basis generators, expressed in the computational basis as

$$\hat{x} : \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \quad \hat{y} : \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix}, \quad \hat{z} : \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}. \tag{1.17}$$

The basis generators of $\mathfrak{su}(2^{N_q})$ consist of all possible $4^{N_q} - 1$ tensor products of the Pauli matrices, omitting the N_q-qubit identity,

$$-i\hat{P}_i = -i\bigotimes_{p=1}^{N_q} \hat{\sigma}_p^{(i)}, \tag{1.18}$$

where $\hat{\sigma}_p^{(i)}$ may be one of the single qubit Pauli matrices $\{\hat{x}, \hat{y}, \hat{z}\}$ along with the single qubit identity acting on the p^{th} qubit. For brevity, $\hat{P}_i$'s are referred to as *Pauli terms* throughout this work. Due to their exponential number, one can not efficiently utilize all generators of $\mathfrak{su}(2^{N_q})$, and hence any practical quantum algorithm must rely on utilizing a sub-exponentially scaling subset thereof to accomplish a task.

In addition to state preparation, one must also be able to measure the state of the quantum register to extract information. This is accomplished by sampling projective measurements of interest in the computational basis $\{0,1\}^{\otimes N_q}$, and the sampling complexity depends on the observables of interest and the quantum algorithm employed. The computational basis is the eigenbasis of $\hat{z}_1 \otimes \ldots \otimes \hat{z}_{N_q}$, and hence alongside unitary generators, it is also natural to express observable quantities in the operator basis of Eq.(1.18). A general

Hermitian operator may be expressed in the qubit algebra as

$$\hat{O} = \sum_i c_i \hat{P}_i, \tag{1.19}$$

where c_i's are real coefficients. Since $\{-i\hat{P}_i\}_{i=1}^{4^{N_q}-1}$ form a complete basis for N_q-qubit anti-Hermitian operators, it is straightforward that any N_q-qubit Hermitian quantity has the form of Eq.(1.19). Since only the $\hat{P}_i$ which consist strictly of $\hat{z}$ operators are directly measurable on the quantum device, $\hat{P}_i$ which are non-diagonal in the computational basis must be diagonalized using additional unitary rotations before measurement.[14–17]

In the following section, we describe how to faithfully encode states and operators relevant to the electronic structure problem to the qubit space.

1.3 Fermion to qubit mappings

The first step of applying quantum computers in solving the electronic structure problem is an encoding from fermionic Fock space objects to those defined on qubit space. Such encodings are referred to as fermion-to-qubit mappings. Noting the similarity of Eq.(1.14) and Eq.(1.15), a Fock occupation vector defined over M spin orbitals is straightforwardly encoded as a $M = N_q$ qubit computational basis state. I.e., there is a one-two-one correspondence between Fock vectors $\left|n_1, \ldots, n_{N_q}\right\rangle$ and computational basis states $\left|z_1, \ldots, z_{N_q}\right\rangle$. The most trivial isomorphism between the two is to set $n_p = z_p$, i.e. qubit p encodes occupation of spin orbital χ_p while in the $|1\rangle$ state, and virtual χ_p while in the $|0\rangle$ state. Such a local encoding of spin-orbital information to qubits is utilized in the Jordan-Wigner (JW) fermion-to-qubit map.[18, 19] Under the JW map, an arbitrary encoded N_e-electron state is represented on a qubit register in the span of computational basis states of Hamming weight N_e, i.e.,

$$|\Psi_{\mathrm{qc}}\rangle = \sum_{\substack{\boldsymbol{z}\in\{0,1\}^{\otimes N_q} \\ \sum_p z_p = N_e}} C_{\boldsymbol{z}} \left|z_1, \ldots, z_{N_q}\right\rangle. \tag{1.20}$$

There exist other non-local mappings of occupancy information, such as the parity and Bravyi-Kitaev (BK) encodings.[20–22] For sake of demonstration, we complete the description of the JW map below. In solving the electronic structure problem using the quantum computer, it is paramount that the electronic Hamiltonian is mapped to a qubit-space observable in an isospectral fashion. It is sufficient to consider a representation of the fermionic creation and annihilation operators in the basis of Pauli terms $\hat{P}_i$'s which conserve the canonical anticommutation relations Eq.(1.9). In the JW encoding, this is accomplished by

defining the equivalencies

$$\hat{a}_p^\dagger = \frac{1}{2}\left(\hat{x}_p - i\hat{y}_p\right) \otimes \hat{z}_{p-1} \ldots \hat{z}_1 \tag{1.21}$$

$$\hat{a}_p = \frac{1}{2}\left(\hat{x}_p + i\hat{y}_p\right) \otimes \hat{z}_{p-1} \ldots \hat{z}_1. \tag{1.22}$$

The electronic Hamiltonian Eq.(1.11) is mapped to qubit-space by a substitution of the fermionic mode operators to elements of the Pauli algebra according to Eq.(1.21) and Eq.(1.22), resulting in the form

$$\hat{H} = \sum_i \eta_i \hat{P}_i, \tag{1.23}$$

where the information carried by the h_{pq} and g_{pqrs} integrals of Eq.(1.11) has been encoded into the Pauli term coefficients η_i. Since the electronic Hamilonian is a two-body operator, it's qubit image Eq.(1.23) consists of $O(N_q^4)$ Pauli terms. Any other fermionic operator, such as symmetries, are encoded in the same fashion. Since the qubit-space $\hat{H}$ possesses identical eigenvalue spectrum to the initial fermionic-space Hamiltonian, it can readily be applied in quantum algorithms for performing ground state energy estimation. In the following section, we provide an overview of one well-explored quantum algorithm for accomplishing this using current and near-term quantum devices.

1.4 The variational quantum eigensolver

There exist many proposals for employing quantum computers in solving the electronic structure problem. One robust approach to determining electronic eigenenergies is the quantum phase estimation (QPE) algorithm,[23] which directly encodes eigenenergies in the readout of the quantum register. However, it utilizes repeated application of the qubit-mapped electronic time evolution operator, $e^{-it\hat{H}}$, leading to circuit decompositions beyond what is implementable on the bleeding-edge quantum hardware devices at the time of writing. Furthermore, the success of measuring the ground state energy is proportional to the squared overlap of the input trial state and the target FCI state, necessitating a qualitatively accurate state preparation routine beforehand. Many variants of the QPE algorithm have been proposed with the goal of being more amenable to the early fault tolerant quantum computing paradigm.[24–27] In order to address the electronic structure problem on current and near-term quantum devices, the variational quantum eigensolver (VQE) was devised, utilizing both quantum and classical compute resources. The VQE is given a parameterized *unitary ansatz*, $\hat{U}(\boldsymbol{\tau})$, where $\boldsymbol{\tau}$ is a set of variational parameters, and seeks to minimize the trial energy through the Rayleigh-Ritz variational principle,

$$\min_{\boldsymbol{\tau}} E(\boldsymbol{\tau}) = \min_{\boldsymbol{\tau}} \left\langle \phi_0 \right| \hat{U}^\dagger(\boldsymbol{\tau}) \hat{H} \hat{U}(\boldsymbol{\tau}) \left| \phi_0 \right\rangle, \tag{1.24}$$

where ϕ_0 is a reference computational basis state. The trial wavefunction is obtained by the current-step unitary rotation, $|\Psi(\boldsymbol{\tau})\rangle = \hat{U}(\boldsymbol{\tau})\,|\phi_0\rangle$. At the current point in parameter space, $\boldsymbol{\tau}$, the trial energy $E(\boldsymbol{\tau})$ is obtained by estimating expectation value $\langle\Psi(\boldsymbol{\tau})|\,\hat{H}\,|\Psi(\boldsymbol{\tau})\rangle$ on the quantum device. The expectation value is sampled by decomposing the qubit-mapped Hamiltonian Eq.(1.23) into simultaneously measurable parts,

$$\hat{H} = \sum_j \hat{O}_j, \tag{1.25}$$

where one has knowledge on how to diagonalize the individual $\hat{O}_j$ in the computational basis. The expectation values $\langle\Psi(\boldsymbol{\tau})|\,\hat{O}_j\,|\Psi(\boldsymbol{\tau})\rangle$ are then constructed by repeated state-preparation followed by measurements, until convergence in the energy estimator is observed, or a measurement shot allocation has been reached. There exist many proposals for obtaining the measurable parts $\{\hat{O}_j\}$, such as partitioning into mutually commuting subsets.[15, 28]

Once the current trial energy has been estimated to sufficient accuracy, it is used as the input to a classical optimizer to suggest an update in parameter space, $\boldsymbol{\tau} \rightarrow \boldsymbol{\tau}^{'}$. If energy gradients are accessible, one can use any standard gradient-based optimizer, such as the Broyden–Fletcher–Goldfarb–Shanno (BFGS) algorithm.[29] Alternatively, gradient free approaches such as constrained optimization by linear approximation (COBYLA) or Nelder Mead can be utilized.[30, 31] There also exist optimization techniques designed specifically for application in variational quantum algorithms, such as quantum natural gradient descent,[32] and techniques based on simultaneous perturbation stochastic approximation.[33]

The VQE algorithm repeats the process of quantum expectation value estimation of $E(\boldsymbol{\tau})$ and classical updating of $\boldsymbol{\tau}$ until a convergence criterion is met, such as $|E(\boldsymbol{\tau}) - E(\boldsymbol{\tau}^{'})|$ dropping below a predefined threshold, or vanishing energy gradient norm. The optimized $E(\boldsymbol{\tau})$ is taken to be the ground state energy estimate for $\hat{H}$. The optimized energy estimate depends on the unitary ansatz $\hat{U}(\boldsymbol{\tau})$ employed. Due to the exponentially scaling number of generators in $\mathfrak{su}(2^{N_q})$, one must fundamentally resort to a subset of generators believed to adequately describe the ground state wavefunction by action of their generated unitary operations. Further, the current technological constraints of existing and near-term hardware necessitate a unitary ansatz which is frugal in gate counts. To this end, a plethora of heuristic unitary ansätze have been proposed. Many proposals have been made which start in the fermionic algebra, borrowing similar intuitions from classical electronic structure techniques. An early and extensively investigated approach is the unitary coupled cluster (UCC) ansatz,[14, 34–39] where one starts with the exponential of an antihermitian analogue of the CC cluster operator,

$$e^{\hat{T}-\hat{T}^\dagger}, \tag{1.26}$$

where the cluster operator $\hat{T}$ is

$$\hat{T} = \sum_{ia} t_i^a \hat{a}_a^\dagger \hat{a}_i + \sum_{ijab} t_{ij}^{ab} \hat{a}_a^\dagger \hat{a}_b^\dagger \hat{a}_i \hat{a}_j + \ldots, \tag{1.27}$$

and $t_{ij\ldots}^{ab\cdots}$ are variational amplitudes. In principle, Eq.(1.27) contains up to N_e-level excitations. However, to avoid prohibitively deep circuits, $\hat{T}$ is often truncated at the singles (S) and doubles (D) level, yielding the UCCSD ansatz in analogy to CCSD. However, unitary $e^{\hat{T}-\hat{T}^\dagger}$ can not be implemented directly, it must be mapped to qubit-space in the same manner as $\hat{H}$. Furthermore, direct implementation of qubit-mapped $e^{\hat{T}-\hat{T}^\dagger}$ as a series of gates is not possible due to general non-commutativity of $\hat{P}_i$ and $\hat{P}_j$ arising in the qubit image of $\hat{T} - \hat{T}^\dagger$. This leads to the necessity of utilizing a Trotter approximation of $e^{\hat{T}-\hat{T}^\dagger}$ to arrive at a product formula of singly exponentiated $e^{-i\tau_i \hat{P}_i}$.*[34] Unfortunately, even low-order Trotterizations of the UCCSD unitary can lead to gate counts outside the scope of current day implementation. While UCCSD is only one approach relying on chemical intuition to heuristically select a polynomial subset of important $\mathfrak{su}(2^{N_q})$ generators, many other fixed ansätze originating from the fermionic picture suffer from deep circuit demands. In a separate direction, there exist "hardware-efficient" ansätze, which do not rely on fermionic intutions, but are rather designed from the standpoint of minimizing circuit depths. Such ansätze involve a fixed set of one-qubit rotations and two-qubit entangling gates, leading to much shallower circuits. However, such hardware-efficient approaches can lead to significant symmetry violations, and can possess serious issues with their optimization.[43, 44]

In the following section, we describe a class of unitary ansätze which are not fixed a priori, but rather constructed on-the-fly to parameterize the ground state in as shallow circuits as possible.

1.5 Adaptive ansätze

For the VQE to provide sufficiently accurate ground state energy estimates, the unitary ansatz $\hat{U}(\boldsymbol{\tau})$ employed must be sufficiently flexible in parameterizing the true ground state. Within the near-term setting, it is paramount that a sufficiently close parameterization of the ground state is achieved with a reasonably shallow circuit. In other words, a practical unitary ansatz must rotate a starting computational basis state sufficiently close to the qubit-encoded electronic ground state in a non-prohibitive number of circuit operations. This requirement has lead to a broad investigation of the so-called *adaptive* ansätze. Rather than having a fixed circuit parameterization, an adaptive ansatz features the selection of unitary generators $\hat{G}_\alpha$ from a predefined operator pool, $\mathcal{P} = \{\hat{G}_\alpha\}$. The ansatz is then "adaptively" constructed iteratively, where at each step, a set of $N_g \geq 1$ generators is

*The casting to an approximate product formula also leads to the so-called "ordering problem",[40] where the order of non-commuting singly exponentiated generators can play an important role in the quality of the ansatz. For instance, see References 41 and 42.

selected to be included in the ansatz based on some importance measure. Following the selection of $\{\hat{G}_\alpha^{(K)}\}_{\alpha=1}^{N_g} \subset \mathcal{P}$, the trial state of the K^{th} iteration is given by

$$|\Psi_K(\boldsymbol{\tau}_1, \ldots \boldsymbol{\tau}_K)\rangle = \hat{U}_K(\boldsymbol{\tau}_K) \, |\Psi_{K-1}(\boldsymbol{\tau}_1, \ldots, \boldsymbol{\tau}_{K-1})\rangle \,, \tag{1.28}$$

where

$$\hat{U}_K(\boldsymbol{\tau}_K) = \prod_{\alpha=1}^{N_g} \exp\Big(-i\tau_\alpha^{(K)} \hat{G}_\alpha^{(K)}\Big), \tag{1.29}$$

and $\boldsymbol{\tau}_K = (\tau_1^{(K)}, \ldots, \tau_{N_g}^{(K)})$ are real variational parameters. The starting state $|\Psi_1\rangle$ is often taken to be a trivially prepared reference, such as the computational basis state encoding the Hartree-Fock occupation vector. The K^{th} energy estimate is given via the minimization of energy with respect to the newly introduced parameters $\boldsymbol{\tau}_K$, with the previously introduced parameters fixed at their optimal values,[†]

$$E_K = \min_{\boldsymbol{\tau}_K} \langle \Psi_K(\boldsymbol{\tau}_1, \ldots, \boldsymbol{\tau}_K)| \, \hat{H} \, |\Psi_K(\boldsymbol{\tau}_1, \ldots, \boldsymbol{\tau}_K)\rangle \,. \tag{1.30}$$

The procedure of generator selection, optimization, and updating of trial state is then repeated until a convergence criterion is met, for instance, if $E_{K-1} - E_K$ falls below a predefined threshold. The defining of pool $\mathcal{P}$, and importance measure used in selection of $\{\hat{G}_\alpha^{(K)}\}_{\alpha=1}^{N_g}$ play crucial roles in the ability of an adaptive ansatz to efficiently parameterize the ground state.

In the context of iteratively parameterizing the electronic ground state, a very reasonable ideal is to minimize the current-iteration energy E_K as much as possible, i.e., to maximize $E_{K-1} - E_K$. Motivated by maximizing the energy lowering per iteration, a widely explored heuristic importance measure for selecting generators from $\mathcal{P}$ is to utilize energy derivatives. Given a generator $\hat{G}_\alpha \in \mathcal{P}$, the derivative of the energy with respect to its associated amplitude τ_α at zero is

$$\left.\frac{\partial E^{(K)}}{\partial \tau_\alpha}\right|_{\tau_\alpha=0} = \frac{\partial}{\partial \tau_\alpha} \langle \Psi_K| \, e^{i\tau_\alpha \hat{G}_\alpha} \hat{H} e^{-i\tau_\alpha \hat{G}_\alpha} \, |\Psi_K\rangle \, \Big|_{\tau_\alpha=0} \tag{1.31}$$

$$= -i \, \langle \Psi_K| \, [\hat{H}^{(K)}, \hat{G}_\alpha] \, |\Psi_K\rangle \,. \tag{1.32}$$

The absolute magnitude of Eq.(1.31) is a direct measure of energy sensitivity to tuning on of $e^{-i\tau_\alpha \hat{G}_\alpha}$ applied to current reference $|\Psi_K\rangle$. Hence, to select the N_g generators to enter the current step unitary, all $G_\alpha \in \mathcal{P}$ are ranked by the absolute values of Eq.(1.31), and the N_g generators of highest values define $\hat{U}_K(\boldsymbol{\tau}_K)$, Eq.(1.29).

[†]One can also consider the reoptimization of all previously parameters $\{\tau_\alpha\}_{\alpha=1}^{K-1}$ simultaneously with optimization of τ_K. While the relaxation of previously introduced parameters in the presence of $e^{-i\tau_K \hat{G}_K}$ can potentially provide a lower E_K, it comes at the cost of a considerably more challenging optimization.

Proposals for operator pools have included generators originating from fermionic operators, such as the excitation/de-excitations defining the UCCSD cluster operator, leading to pool $\mathcal{P}_{SD} = \{\hat{a}_a^\dagger \hat{a}_i - \text{h.c.}\}_{ia} \bigcup \{\hat{a}_a^\dagger \hat{a}_b^\dagger \hat{a}_i \hat{a}_j - \text{h.c.}\}_{ijab}$, such as used in the ADAPT-VQE method.[45] Upon a mapping to qubit-space, such pool elements are linear combinations of Pauli terms. In the JW representation, singles and doubles are represented as

$$\hat{a}_a^\dagger \hat{a}_i - \text{h.c.} \xrightarrow{\text{JW}} \hat{G}_{(ia)} = \frac{i}{2} \bigotimes_{k=i+1}^{a-1} \hat{z}_k \left(\hat{y}_i \hat{x}_a - \hat{x}_i \hat{y}_a\right) \tag{1.33}$$

$$\begin{aligned}\hat{a}_a^\dagger \hat{a}_b^\dagger \hat{a}_i \hat{a}_j - \text{h.c.} \xrightarrow{\text{JW}} \hat{G}_{(ijab)} = \frac{i}{8} \bigotimes_{k=i+1}^{j-1} \hat{z}_k \bigotimes_{l=a+1}^{b-1} \hat{z}_l (&\hat{x}_i \hat{x}_j \hat{y}_a \hat{x}_b + \hat{y}_i \hat{x}_j \hat{y}_a \hat{y}_b + \hat{x}_i \hat{y}_j \hat{y}_a \hat{y}_b + \hat{x}_i \hat{x}_j \hat{x}_a \hat{y}_b \\ &- \hat{y}_i \hat{x}_j \hat{x}_a \hat{x}_b - \hat{x}_i \hat{y}_j \hat{x}_a \hat{x}_b - \hat{y}_i \hat{y}_j \hat{y}_a \hat{x}_b - \hat{y}_i \hat{y}_j \hat{x}_a \hat{y}_b).\end{aligned} \tag{1.34}$$

Fortunately, the Pauli terms in $\hat{G}_{(ia)}$ and $\hat{G}_{(ijab)}$ mutually commute amongst themselves, and hence unitary rotations generated $\hat{G}_{(ia)}$ and $\hat{G}_{(ijab)}$ may be performed on the circuit without the need for Trotterization. However, the qubit-mapped double excitation/de-excitations in Eq.(1.34) leading to eight distinct Pauli rotations can put a considerable limitation on the number of generators N_g entering the current step unitary and the maximum K before circuit description of $|\Psi_K\rangle$ becomes regretfully deep. Furthermore, the cardinalities of $\{\hat{G}_{(ia)}\}_{ia}$ and $\{\hat{G}_{(ijab)}\}_{ijab}$ scale as $O(N_q^2)$ and $O(N_q^4)$ respectively, and hence ranking the elements of $\mathcal{P}$ by importance leads to $O(N_q^4)$ unique quantum gradient estimations in the form of Eq.(1.31). Many other adaptive schemes have been devised to reduce the onset of circuit depth by use of more primitive generators,[46, 47] and to circumvent quantum gradient estimation overhead by use of smaller pool sizes.[48]

All adaptive schemes are designed with the motivation of accurately solving for the electronic ground state using shallow circuits, by iteratively rotating the trial state using unitary operations deemed most important in variational energy lowering. However, explicitly ranking the pool of generators by energy derivatives can lead to a significant increase in the measurement complexity of a single iteration, even if gradient-free approaches are used for optimization Eq.(1.30). Such an overhead should not be overlooked, as measuring the energy expectation value itself is considered the major computational bottleneck of the VQE.[49] Estimating the energy expectation value to precision ϵ requires a number of measurements

$$N_{\text{meas.}}(\epsilon) = \frac{P}{\epsilon^2}, \tag{1.35}$$

where P is a proportionality term, depending on the system Hamiltonian $\hat{H}$, the trial state $|\Psi_K\rangle$, and the energy estimator measurement strategy utilized.[14, 49–52] For ab initio quantum chemistry applications, the target precision for obtained electronic energies is the so-called *chemical accuracy*, an additive error from the exact solution within

1kcal/mol $\approx$ 1.6 mHartrees.[‡] It is clearly undesirable for the final energy estimate to have a sampling error ϵ above this threshold. The large number of measurements necessary to reach such a high precision has been deemed a substantial roadblock in utilizing the VQE in demonstrating quantum advantage,[49] and the situation is worsened if one also has to estimate the energy derivatives of all $\hat{G}_\alpha \in \mathcal{P}$ to reasonable precision.

In the following chapter, we introduce the iterative qubit coupled cluster (iQCC) algorithm, which has many similarities to adaptive VQE schemes as described here, while also possessing important differences. These differences allow us to utilize a generator pool which is totally unrestricted: the basis generators of $\mathfrak{su}(2^{N_q})$ are taken as $\mathcal{P}$. Remarkably, the iQCC algorithm utilizes a fully classical algorithm to constructively obtain the elements of high importance in $\mathcal{P}$ in an efficient manner, avoiding any need for quantum gradient estimations. The algorithm relies on structure of the Hamiltonian itself to identify which generators will have a large impact on variational energy lowering. Further, by adopting a "Heisenberg picture" approach to growing the adaptive ansatz, one can maintain constant circuit depth, at the expense of growing complexity of the Hamiltonian. In unison, these aspects allow us to implement iQCC as a fully classical quantum-inspired algorithm, or as a classical state-preparation routine for QPE, avoiding the VQE measurement roadblock altogether. We describe such aspects in detail below.

[‡]For a quantum chemical calculation to be predictively accurate, it is desirable that errors are within experimental accuracy. John Pople states in his Nobel Lecture (1999): "For energies, such as heats of formation or ionization potentials, a global accuracy of 1 kcal/mol would be appropriate."[53] For certain predictive applications, sub-chemical accuracy may be desirable. For instance, "spectroscopic accuracy" often refers to vibrational frequencies within $\pm 1\text{cm}^{-1}$ of experiment, generally translating to a desired energy precision closer to a microHartree.[54]

Chapter 2

The iterative qubit coupled cluster algorithm

Within this chapter, we describe an adaptive approach to estimating ground state electronic energies, referred to as the *iterative qubit coupled cluster* (iQCC) method. The iQCC algorithm, while a heuristic method, has shown robust numerical convergences towards FCI solutions for challenging instances of the electronic structure problem.* Unlike many other ab initio electronic structure techniques, it is formally derived in the algebra of qubits, rather than being derived in the fermionic algebra and mapped to qubit algebra for application on a quantum device. The philosophy of working directly in qubit space was originally put forth due to the limited circuit depths of current and near-term quantum hardware. The mapping of fermionic excitation/de-excitations, e.g. Eqs. (1.33) and (1.34) lead to linear combinations of Pauli terms $\hat{P}_\alpha$, i.e., multiple basis generators of $\mathfrak{su}(2^{N_q})$. In fact, the number of such basis generators spanned in the qubit image of fermionic excitation/de-excitation operators is exponential with respect to the fermionic excitation rank,[55] leading to high circuit demand for triple excitation/de-excitations and higher. The iQCC algorithm instead selects individual $\hat{P}_\alpha$'s of high energy derivative to be included in the unitary ansatz, directly informed by the Hamiltonian. However, for sufficiently difficult instances of the electronic structure problem, one may still require a sequence of an exceedingly high number of basis generator rotations $\exp\left(-i\tau_\alpha \hat{P}_\alpha\right)$ to bring an initial reference state within chemical accuracy of the target ground state. To circumvent the circuit depth constraint on the unitary ansatz, the iQCC algorithm utilizes a strategy based on the Heisenberg picture, unitarily transforming the qubit-mapped electronic Hamiltonian $\hat{H}$ with the previously obtained $\exp\left(-i\tau_\alpha \hat{P}_\alpha\right)$. For example, if one has a parameterization of the target state as $\hat{U}_A \hat{U}_B \left|\phi_0\right\rangle$, where $\hat{U}_A$ and $\hat{U}_B$ are unitary transformations, acting on a reference state

*At time of writing, the largest published iQCC simulations can be found in Reference 5, where transition energies were calculated for Ir(III) complexes within active spaces of 36 electrons in 36 orbitals. Such calculations utilized iQCC as a classical algorithm, where the corresponding quantum implementation would be defined over 72 qubits.

$|\phi_0\rangle$, one can access the target energy by estimating e.g., $\langle\phi_0|\hat{U}_B^\dagger\hat{H}_A\hat{U}_B|\phi_0\rangle$, where $\hat{H}_A = \hat{U}_A^\dagger\hat{H}\hat{U}_A$, where only $\hat{U}_B$ is left for circuit implementation. Having outlined a few of the motivations for the iQCC algorithm, we concretely outline the steps of the procedure in the following section, and further expand on the individual steps in the subsequent sections.

2.1 Overview

The iQCC method is an adaptive variational algorithm utilizing effective Hamiltonian transformations. The method involves an iterative similarity transforming of the qubit-space Hamiltonian $\hat{H}$ with parameterized unitary operations. Herein, a summary of the essential parts of the method is given, omitting many of the technical details. The following sections of this chapter provide further insight into the computational aspects involved in the steps of the iQCC procedure.

The iQCC algorithm consists of iterating the process of: selecting unitary generators based on the Hamiltonian, optimizing their associated variational amplitudes, and subsequent unitary transformation of a current-step effective Hamiltonian. Such process defines a step of the iQCC procedure. An iQCC step begins with the selection of $N_g \geq 1$ Pauli terms, $\{\hat{P}_\alpha^{(K)}\}_{\alpha=1}^{N_g}$, which are utilized to parameterize the current K^{th} iteration iQCC unitary,

$$\hat{U}_K(\boldsymbol{\tau}_K) = \prod_{\alpha=1}^{N_g} \exp\Big(-i\tau_\alpha^{(K)}\hat{P}_\alpha^{(K)}/2\Big), \tag{2.1}$$

where $\{\tau_\alpha^{(K)}\}_{\alpha=1}^{N_g}$ are optimizable amplitudes. The Pauli terms $\{\hat{P}_\alpha\}_{\alpha=1}^{N_g}$ are selected for inclusion in Eq.(2.1) based on a *screening* procedure utilizing energy gradients. Such gradients essentially capture the magnitude of coupling between the reference state and the excited configuration generated by $\hat{P}_\alpha$, see Section 2.2 for an in depth description. Following the selection of unitary generators, the iQCC procedure aims to variationally lower the current electronic energy

$$\min_{\boldsymbol{\tau}_K} E_K(\boldsymbol{\tau}) = \min_{\boldsymbol{\tau}_K} \langle\phi_0|\hat{U}_K^\dagger(\boldsymbol{\tau}_K)\hat{H}^{(K)}\hat{U}_K(\boldsymbol{\tau}_K)|\phi_0\rangle \tag{2.2}$$

where $|\phi_0\rangle$ denotes a reference computational basis state,[†] i.e., a qubit-mapped electronic configuration in the orbital basis of the original fermion-to-qubit mapped electronic Hamil-

[†]Alternatively, a so-called *qubit mean-field* reference state can be used,[1, 46] which is parameterized by $2N_q$ single qubit Bloch angles, and lives in the manifold of all single qubit product states,

$$|\phi_0(\boldsymbol{\theta},\boldsymbol{\psi})\rangle = \bigotimes_{p=1}^{N_q}\Big(\cos(\theta_p)|0\rangle + e^{i\psi_p}\sin(\theta_p)|1\rangle\Big). \tag{2.3}$$

While re-optimization of the Bloch angles at every iteration in principle can offer lower trial energies, it poses a more difficult optimization, and can lead to exacerbated symmetry violations.[46, 56, 57] For simplicity of discussion throughout, we assume a computational basis reference state, e.g., the Hartree-Fock state

tonian, $\hat{H}^{(1)}$. A standard choice of such computational basis state is the qubit-mapped Hartree-Fock state. Unlike many other adaptive schemes employing variational quantum circuits, the iQCC method involves a similarity transformation, or *dressing*, of the current iteration Hamiltonian, rather than appending the newly introduced unitary to a quantum circuit ansatz. Such dressing yields the updated effective Hamiltonian for the next iteration, and is the final step of current iteration K,

$$\hat{H}^{(K+1)} = \hat{U}_K^\dagger(\boldsymbol{\tau}_K^*)\hat{H}^{(K)}\hat{U}_K(\boldsymbol{\tau}_K^*), \tag{2.4}$$

where $\boldsymbol{\tau}_K^*$ are the optimal amplitudes obtained through minimization Eq.(2.2). This transformation is performed by recursively applying single Pauli exponential dressing for each factor in $\hat{U}_K(\boldsymbol{\tau}_K^*) = \prod_{\alpha=1}^{N_g} \exp\Big(-i\tau_\alpha^{(K)}\hat{P}_\alpha^{(K)}/2\Big)$. Such a single Pauli exponential dressing is given by

$$e^{i\tau_\alpha\hat{P}_\alpha/2}\hat{H}e^{-i\tau_\alpha\hat{P}_\alpha/2} = \hat{H} - \frac{i}{2}\sin(\tau_\alpha)[\hat{H},\hat{P}_\alpha] + \frac{1}{2}\left(1-\cos(\tau_\alpha)\right)\Big(\hat{P}_\alpha\hat{H}\hat{P}_\alpha - \hat{H}\Big) \tag{2.5}$$

The $(K+1)^{\text{th}}$ iteration then utilizes effective $\hat{H}^{(K+1)}$ in generator selection and optimization of E_{K+1}. The process of *selection*, *optimization*, and *transformation* define a single iteration of the iQCC procedure, and are repeated until a convergence criterion is achieved. The convergence criterion can be taken to be e.g., $|\Delta E| = |E_K - E_{K-1}|$ or highest observed value of g_α at iteration K falling below a threshold. The final ground state energy estimate is then E_K for convergence criterion met at iteration K.

The number of terms in effective $\hat{H}^{(K)}$ *formally* scales as $M_K = O(M_1(3/2)^{N_g K})$, and hence the iQCC procedure formally scales exponentially for any $N_g \geq 1$. However, large scale numerical calculations[5] have shown that, under reasonable truncation of Hamiltonian terms, M_K is often far below the theoretical exponential scaling.[‡] However, as presented in Section 2.5, the growing number of terms in the iQCC effective Hamiltonians remains as the main computational bottleneck.

For clarity, we provide a summary of the three steps defining an iQCC iteration below.

1. Selection: Select $N_g \geq 1$ Pauli terms for inclusion as generators in current step unitary $\hat{U}(\boldsymbol{\tau})$ [Eq.(2.1)] by highest values of energy gradient magnitude g_α based on the current step Hamiltonian [Section 2.2].

2. Optimization: Minimize current step energy with respect to the N_g variational parameters $\boldsymbol{\tau}$ Eq.(2.2), yielding the current step energy estimate and optimized parameter values $\boldsymbol{\tau}^*$ [Section 2.3].

3. Transformation: Obtain the next-step effective Hamiltonian via similarity transformation of current step effective Hamiltonian by $\hat{U}(\boldsymbol{\tau}^*)$ as Eq.(2.24). [Section 2.4]

[‡]Reference 5 reports that at the late-stage iterations, the effective Hamiltonian growth could be fit to a low scaling polynomial, $\sim M^{1.6}$.

By construction, the iQCC procedure is rather flexible in implementation:

1. As a NISQ variational quantum algorithm, where unitary Eq.(2.1) is performed as a quantum circuit, and optimization Eq.(2.2) is carried out as a VQE subroutine.

2. As a "quantum-inspired" fully classical method, where circuit optimization is replaced by a classical classical energy evaluation feasible for modest N_g (see Section 2.3 for more details).

3. As a fully classical state-preparation routine for QPE. With the iQCC procedure converging towards the ground state, the total iQCC unitary at iteration K, $\hat{U}_{tot} = \prod_{i=1}^{K} \hat{U}_i(\boldsymbol{\tau}_i^*)$ achieves higher overlap with the true ground state $|\Psi_0\rangle$ than reference $|\phi_0\rangle$, i.e., $\left|\langle\Psi_0|\,\hat{U}_{tot}\,|\phi_0\rangle\right|^2 > |\langle\Psi_0|\phi_0\rangle|^2$, and hence can be used to improve the success probability of measuring the ground state in QPE.

2.2 Selection of generators

In order to compactly parameterize the iQCC unitary $\hat{U}_{tot} = \prod_{i=1}^{K} \hat{U}_i(\boldsymbol{\tau}_i)$ which achieves $\hat{U}_{tot}\,|\phi_0\rangle \approx |\Psi_0\rangle$, it is paramount that the basis generators $\hat{P}_\alpha$ entering each $\hat{U}_i$ are selected in a judicious and efficient manner. The number of basis generators of $\mathfrak{su}(2^{N_q})$ is $4^{N_q} - 1$, and hence brute force ranking of the individual basis generators is highly unfeasible. As it turns out, one can devise a fully classical algorithm to *constructively* obtain generators $\hat{P}_\alpha$ of non-zero energy derivative, which is efficient in N_q and the number of terms in the Hamiltonian, M. Herein, we review the formalism for classically identifying (*screening*) Pauli terms from the unrestricted pool $\{\hat{P}_i\}_{i=1}^{4^N-1}$ which possess non-zero energy gradient, given the generated unitary acts on a single Slater determinant reference. The key to efficiently performing such an unrestricted screening is the formulation of the "direct interaction space" (DIS), a set coming directly from inspection of the Hamiltonian, guiding the construction of high gradient Pauli terms, which can be computed efficiently classically.[1]

Considering a computational basis state reference $|\phi_0\rangle$, i.e., a single Slater determinant, the iQCC energy cost function to minimize is

$$E(\tau_\alpha) = \langle\phi_0|\, e^{i\tau_\alpha \hat{P}_\alpha/2} \hat{H} e^{-i\tau_\alpha \hat{P}_\alpha/2}\,|\phi_0\rangle\,, \tag{2.6}$$

where $\hat{H}$ is assumed to be the qubit-mapped Hamiltonian at an arbitrary iteration. We wish to identify $\hat{P}_\alpha$ which have highest impact on lowering the energy, i.e., $\hat{P}_\alpha$ with largest expected $E(0) - E(\tau_\alpha^*)$, where τ_α^* is the minimizing amplitude for $E(\tau_\alpha)$. As a guiding proxy

to expected energy lowering, we utilize the energy derivative evaluated at $\tau_\alpha = 0$,

$$\frac{\partial E}{\partial \tau_\alpha}\Big|_{\tau_\alpha=0} = \frac{\partial}{\partial \tau_\alpha} \langle \phi_0 | e^{i\tau_\alpha \hat{P}_\alpha/2} \hat{H}^{(K)} e^{-i\tau_\alpha \hat{P}_\alpha/2} | \phi_0 \rangle \tag{2.7}$$

$$= -\frac{i}{2} \langle \phi_0 | [\hat{H}, \hat{P}_\alpha] | \phi_0 \rangle \tag{2.8}$$

$$= \mathrm{Im}(\langle \phi_0 | \hat{H} \hat{P}_\alpha | \phi_0 \rangle). \tag{2.9}$$

The last equality is met due to the Hermitian property of $\hat{H}$ and $\hat{P}_\alpha$. We refer to the absolute magnitude of Eq.(2.7) as g_α, that is,

$$g_\alpha = \left| \frac{\partial E}{\partial \tau_\alpha}\Big|_{\tau_\alpha=0} \right|. \tag{2.10}$$

We then consider the qubit-mapped Hamiltonian in the following form:

$$\hat{H} = \sum_k \eta_k \hat{P}_k \tag{2.11}$$

$$= \sum_i \left(\sum_j \eta_j^{(i)} \hat{Z}_j^{(i)} \right) \hat{X}_i, \tag{2.12}$$

where such representation is always valid due to the fact that any arbitrary Pauli term may be expressed as a product of strictly $\hat{z}$ Pauli operators, $\hat{Z}_j^{(i)}$, multiplying a product of strictly $\hat{x}$ Pauli operators, $\hat{X}_i$, up to a phase of $\{1, -1\}$,[§] which we can absorb into the coefficient $\eta_j^{(i)}$. Inserting such a representation of $\hat{H}$ into Eq.(2.10), we obtain

$$g_\alpha = \left| \mathrm{Im} \left(\sum_i \langle \phi_0 | \left(\sum_j \eta_j^{(i)} \hat{Z}_j^{(i)} \right) \hat{X}_i \hat{P}_\alpha | \phi_0 \rangle \right) \right|. \tag{2.13}$$

The summation $\sum_j \eta_j^{(i)} \hat{Z}_j^{(i)}$ can be seen acting directly on the reference dual state $\langle \phi_0 |$, which is the adjoint of a mutual eigenstate of the strictly Pauli $\hat{z}$ products, giving

$$g_\alpha = \left| \mathrm{Im} \left(\sum_i \left(\sum_j \eta_j^{(i)} \lambda_j^{(i)} \right) \langle \phi_0 | \hat{X}_i \hat{P}_\alpha | \phi_0 \rangle \right) \right|, \tag{2.14}$$

with $\hat{Z}_j^{(i)} | \phi_0 \rangle = \lambda_j^{(i)} | \phi_0 \rangle$ where $\lambda_j^{(i)} \in \{1, -1\}$. The imaginary component of expectation

[§]Generally speaking, the multiplication of totally arbitrary $\hat{Z}_k$ and $\hat{X}_l$ can also produce imaginary phases $\{i, -i\}$, if they possess odd overlapping support. However, since the Hamiltonian is purely real-valued, it turns out that all of the Pauli terms occurring in qubit-mapped $\hat{H}$ have an even number of $\hat{y}_p$ operations, with real coefficients, and hence the phase incurred by multiplication of $\hat{Z}_j^{(i)}$ and $\hat{X}_i$ in Eq.(2.12) is always real. This property is conserved during the iQCC Hamiltonian dressing.

value $\langle\phi_0|\,\hat{X}_i\hat{P}_\alpha\,|\phi_0\rangle$ must be vanishing unless $\hat{P}_\alpha$ meets certain requirements. Let

$$\hat{X}_i = \prod_{p=1}^{n} \hat{x}_p^{\mu_p^{(i)}}, \tag{2.15}$$

$$\hat{Z}_i = \prod_{p=1}^{n} \hat{z}_p^{\nu_p^{(i)}}, \tag{2.16}$$

where $\mu_p^{(i)}, \nu_p^{(i)} \in \{0,1\}$. Expressing $\hat{P}_\alpha$ as $\theta_\alpha \hat{X}_\alpha \hat{Z}_\alpha$ with $\theta_\alpha = e^{i\pi \sum_{p=1}^{n} \mu_p^{(\alpha)} \nu_p^{(\alpha)}/2}$ (θ_α is introduced to cancel any prefactor introduced in multiplication of factors $\hat{X}_\alpha$ and $\hat{Z}_\alpha$, ensuring the factorization yields a phaseless Pauli term),

$$g_\alpha = \left|\mathrm{Im}\left(\theta_\alpha \sum_i \left(\sum_j \eta_j^{(i)} \lambda_j^{(i)}\right) \langle\phi_0| \prod_{p=1}^{n} \hat{x}_p^{\mu_p^{(i)}}\, \hat{x}_p^{\mu_p^{(\alpha)}}\, \hat{Z}_\alpha\, |\phi_0\rangle\right)\right|, \tag{2.17}$$

$$= \left|\mathrm{Im}\left(\lambda_\alpha \theta_\alpha \sum_i \left(\sum_j \eta_j^{(i)} \lambda_j^{(i)}\right) \langle\phi_0| \prod_{p=1}^{n} \hat{x}_p^{\mu_p^{(i)}}\, \hat{x}_p^{\mu_p^{(\alpha)}}\, |\phi_0\rangle\right)\right|, \tag{2.18}$$

$$= \left|\mathrm{Im}\left(\lambda_\alpha \theta_\alpha \sum_i \left(\sum_j \eta_j^{(i)} \lambda_j^{(i)}\right) \prod_{p=1}^{n} \delta_{\mu_p^{(i)}, \mu_p^{(\alpha)}}\right)\right|, \tag{2.19}$$

where $\hat{Z}_\alpha\,|\phi_0\rangle = \lambda_\alpha\,|\phi_0\rangle$ and $\delta_{i,j}$ is the Kronecker delta function. Since $c_j^{(i)}$ are strictly real, the only quantity with potentially nonzero imaginary part is θ_α, hence

$$g_\alpha = \left|\mathrm{Im}(\theta_\alpha)\lambda_\alpha \sum_i \left(\sum_j \eta_j^{(i)} \lambda_j^{(i)}\right) \prod_{p=1}^{n} \delta_{\mu_p^{(i)}, \mu_p^{(\alpha)}}\right|, \tag{2.20}$$

$$= \left|\mathrm{Im}(\theta_\alpha)\lambda_\alpha \sum_i \left(\sum_j \eta_j^{(i)} \lambda_j^{(i)}\right) \boldsymbol{\delta}_{\vec{\mu}^{(i)}, \vec{\mu}^{(\alpha)}}\right|, \tag{2.21}$$

where $\vec{\mu}^{(j)} \equiv (\mu_1^{(j)}, \ldots \mu_n^{(j)})$, and $\boldsymbol{\delta}_{\vec{i},\vec{j}}$ is the multidimensional Kronecker delta function, i.e. $\boldsymbol{\delta}_{\vec{i},\vec{j}} = 1$ if $\vec{i}$ and $\vec{j}$ are identical vectors, otherwise $\boldsymbol{\delta}_{\vec{i},\vec{j}} = 0$.

Since $\mathrm{Im}(\theta_\alpha) = \sin\left(\pi \sum_{p=1}^{n} \mu_p^{(\alpha)} \nu_p^{(\alpha)}/2\right)$, it is evident that g_α is vanishing unless $\hat{P}_\alpha$ possesses an odd number of Pauli $\hat{y}$ operators. Furthermore, given $\hat{P}_\alpha$ with odd number of Pauli $\hat{y}$'s, g_α is uniquely determined by $\vec{\mu}^{(\alpha)}$. Candidate binary string $\vec{\mu}^{(\alpha)}$ must be equal to one of the binary strings $\vec{\mu}^{(i)}$ associated to some $\hat{X}_i$ in Eq.(2.12) if $\hat{P}_\alpha$ is to have non-zero g_α, evident by Eq.(2.21). We refer to the set of all binary strings associated to $\hat{X}_i$ operators in Eq.(2.12) as $\Omega = \{\vec{\mu}^{(i)}\}_i$. The rules for determining g_α for candidate Pauli term $\hat{P}_\alpha$ are summarized as follows:

1. If $\hat{P}_\alpha$ does not possess an odd number of $\hat{y}$ instances, i.e., if $\vec{\mu}^{(i)} \cdot \vec{\nu}^{(i)} \mod 2 = 0$, then $g_\alpha = 0$.

2. If $\vec{\mu}^{(\alpha)}$ is not identical to any $\vec{\mu}^{(i)}$ arising from the set of products of Pauli x operators in Eq. (2.12), $\{\hat{X}_i\}_i$, i.e. $\boldsymbol{\delta}_{\vec{\mu}^{(i)},\vec{\mu}^{(\alpha)}} = 0 \,\forall\, \vec{\mu}^{(i)} \in \Omega$, then $g_\alpha = 0$.

3. Otherwise, $\boldsymbol{\delta}_{\vec{\mu}^{(i)},\vec{\mu}^{(\alpha)}} = 1$ for some i, then $g_\alpha \equiv g_i = \left|\sum_j \eta_j^{(i)} \lambda_j^{(i)}\right|$.

Therefore, to efficiently screen Pauli terms from the unrestricted pool $\{\hat{P}_\alpha\}_{\alpha=1}^{4^n-1}$ possessing non-zero gradient: obtain the set Ω, and for each $\vec{\mu}^{(i)} \in \Omega$, g_i is computed as $\left|\sum_j \eta_j^{(i)} \lambda_j^{(i)}\right|$. By doing so, one obtains the *direct interaction space* (DIS) of $\hat{H}$, the set

$$\mathcal{D}_H = \left\{ \left(\vec{\mu}^{(i)}, g_i \right) : \vec{\mu}^{(i)} \in \Omega \right\}. \tag{2.22}$$

The $\vec{\mu}^{(i)}$ with highest g_i tells us how to specify the $\hat{X}_i$ part in a highest gradient Pauli term according to Eq.(2.15), however the only constraint on the corresponding $\hat{Z}_i$ part is that it produces an odd number of $\hat{y}$ instances. There is a large degree of freedom in specifying the $\hat{Z}_i$ part to satisfy $\vec{\mu}^{(i)} \cdot \vec{\nu}^{(i)} \mod 2 = 1$. In the standard iQCC formulation [1], a single $\hat{z}$ instance is put on the first qubit index acted on non-trivially by $\hat{X}_i$ to satisfy the odd y parity constraint. In Chapters 3, 4, and 5, the gauge freedom in specifying $\vec{\nu}^{(i)}$'s up to $\vec{\mu}^{(i)} \cdot \vec{\nu}^{(i)} \mod 2 = 1$ is exploited to yield other desirable properties, such as mitigating the number of terms introduced upon dressing. [3] In Appendix A, we discuss the properties of the non-zero gradient Pauli terms in the fermionic picture.

Essentially, the set of unique $\hat{X}_i$ present in the Hamiltonian directly inform us how to construct the iQCC generators of highest energy gradient. This allows us to formulate the DIS Eq.(2.22) efficiently in N_q and the number of terms in $\hat{H}$.

2.3 Energy optimization

Once a set of N_g generators $\{\hat{P}_\alpha^{(K)}\}_{\alpha=1}^{N_g}$ has been selected for the K^{th} iteration, a prescription for ordering them in the current $\hat{U}_K$ [Eq.(2.1)] must be applied, due to their general non-commutativity.[40] This is standardly done by ordering the generated unitary operations $\exp\left(-i\tau_\alpha^{(K)} \hat{P}_\alpha^{(K)}/2\right)$ by the gradient magnitude g_α's, such that the reference state is acted on by the generated unitary operations in descending order of the generator's gradient magnitude. With the current-step unitary ansatz completely specified, the next step is to optimize the associated amplitudes $\boldsymbol{\tau}_K = \{\tau_\alpha^{(K)}\}_{\alpha=1}^{N_g}$. For general N_g, this energy optimization can be done by means of a VQE-style quantum-classical feedback loop, barring circuit depth restrictions of the quantum device. This amounts to performing optimization Eq.(2.2), yielding the optimal $\boldsymbol{\tau}_K^* = \arg\min_{\boldsymbol{\tau}_K} E_K$ and current step optimal trial energy $\min_{\boldsymbol{\tau}_K} E_K$. As mentioned in Section 1.4, such a VQE optimization may be carried out using many different optimization schemes, including gradient based approaches utilizing analytic derivatives.[58–60]

We now discuss the problem of energy optimization if one wishes to perform the iQCC algorithm in a purely classical manner, which is feasible for small N_g. Replacing quantum

expectation estimate of VQE with a classical energy evaluation enables a fully classical implementation of the iQCC procedure. Considering $N_g = 1$, with single generator $\hat{P}^{(K)}$, the energy to be optimized is

$$E(\tau^{(K)}) = \left\langle \hat{H}^{(K)} \right\rangle + \sin\left(\tau^{(K)}\right) \mathrm{Im}\left(\left\langle \hat{H}^{(K)} \hat{P}^{(K)} \right\rangle\right) + \frac{1}{2}\left(1 - \cos\left(\tau^{(K)}\right)\right)\left(\left\langle \hat{P}^{(K)} \hat{H}^{(K)} \hat{P}^{(K)} \right\rangle - \left\langle \hat{H}^{(K)} \right\rangle\right), \tag{2.23}$$

where expectation values are taken with respect to reference $|\phi_0\rangle$. Noting that $\langle \hat{H}^{(K)} \rangle$ and $\mathrm{Im}\langle \hat{H}^{(K)} \hat{P}^{(K)} \rangle$ are respectively the energy of the previous iteration, and the gradient of current step $\hat{P}^{(K)}$, they have already been computed. The only new quantity to compute is the expectation value $\langle \phi_0 | \hat{P}^{(K)} \hat{H}^{(K)} \hat{P}^{(K)} | \phi_0 \rangle$. It can be shown that this operator has the exact same terms in $\hat{H}^{(K)}$, and only possesses sign differences, letting us straightforwardly to obtain this value in $O(M_K)$ time. The optimal $\tau^{(K)}$ is readily obtained by considering the minimization of $E(\tau^{(K)})$ once the expectation values have been classically computed. Note that such classical energy optimization can be extended to $N_g \geq 1$, however it becomes exponentially complex with N_g, with the energy functional becoming considerably more complex with each additional generator included.

2.4 Hamiltonian transformation

The final step of an iQCC iteration is to similarity transform the Hamiltonian with the optimized $\hat{U}(\boldsymbol{\tau}^*) = \prod_{\alpha=1}^{N_g} \exp\left(-i\tau_\alpha^* \hat{P}_\alpha / 2\right)$. Also owing to the fact that an arbitrary Pauli term is involutory ($\hat{P}_\alpha^2 = \hat{1}$), similarity transforming the Hamiltonian with a single exponentiated Pauli term yields closed form expression:

$$e^{i\tau_\alpha^* \hat{P}_\alpha/2} \hat{H} e^{-i\tau_\alpha^* \hat{P}_\alpha/2} = \hat{H} - \frac{i}{2} \sin\tau_\alpha^* [\hat{H}, \hat{P}_\alpha] + \frac{1}{2}(1 - \cos\tau_\alpha^*)\left(\hat{P}_\alpha \hat{H} \hat{P}_\alpha - \hat{H}\right). \tag{2.24}$$

Transformation Eq.(2.24) is classically performed recursively for all N_g singly exponentiated generators of $\hat{U}(\boldsymbol{\tau}^*)$, which yields the next-step effective iQCC Hamiltonian.

The iQCC transformation has many desirable properties. For one, due to the closed form of Eq.(2.24), one obtains the exactly similarity transformed $U^\dagger(\boldsymbol{\tau}^*)\hat{H}U(\boldsymbol{\tau}^*)$, which is isospectral to $\hat{H}$. This is in contrast to many other Hamiltonian processing techniques which employ approximations to the Baker-Campbell-Hausdorff expansion[¶] of the similarity transformed Hamiltonian, due to the use of more complicated non-involutory generators. [61–65] Secondly, since $\hat{U}$ is unitary, the Hermitian property of the Hamiltonian is conserved,

[¶]For similarity transformation $\hat{A}^{-1}\hat{H}\hat{A}$, where $\hat{A} = e^{\hat{X}}$, the Baker-Campbell-Hausdorff (BCH) expansion expresses $\hat{A}^{-1}\hat{H}\hat{A}$ in terms of nested commutators,

$$\hat{A}^{-1}\hat{H}\hat{A} = e^{-\hat{X}}\hat{H}e^{\hat{X}} = \hat{H} + [\hat{X}, \hat{H}] + \frac{1}{2!}[\hat{X}, [\hat{X}, \hat{H}]] + \frac{1}{3!}[\hat{X}, [\hat{X}, [\hat{X}, \hat{H}]]] + \ldots. \tag{2.25}$$

facilitating subsequent standard variational treatment. However, the iterative dressing of the effective Hamiltonian can lead to the accumulation of many new Pauli terms, and is the major computational bottleneck in the iQCC procedure. By inspection of Eq.(2.24), one sees 3^{N_g} distinct operator terms are produced by transforming $\hat{H}$ with N_g generators. However, the number of unique Pauli terms in the transformed Hamiltonian is well-estimated to be $\sim M(3/2)^{N_g}$, for an initial Hamiltonian of M Pauli terms. This is due to the fact that, firstly, $\hat{P}_\alpha \hat{H} \hat{P}_\alpha$ in Eq.(2.24) possess the same set of unique Pauli terms in $\hat{H}$. Secondly, we expect the commutator $[\hat{H}, \hat{P}_\alpha]$ to produce $M/2$ Pauli terms in the average case.[‖] Hence, we expect Eq.(2.24) to produce an effective Hamiltonian with $\frac{3}{2}M$ terms, and $\sim M(3/2)^{N_g}$ terms for N_g dressings. While the Pauli terms of $\hat{H}$ and $\hat{P}_\alpha$ selected for via high energy gradient possess more structure than being uniformly sampled in practice, this average-case scaling holds reasonably accurate for the initial iterations.[1, 2] If one employs a numerical compression of the dressed Hamiltonian terms, i.e., if one uses

$$\hat{H}_\epsilon = \sum_{\substack{i \\ |\eta_i| \geq \epsilon}} \eta_i \hat{P}_i, \tag{2.27}$$

it is seen that the average case scaling $\sim M(3/2)^{N_q}$ can become a drastic overestimate at high N_g counts, even for fairly precise ϵ.[**][5]

2.5 Results

Within this section, we present numerical results for the iQCC method, aimed at illustrating its advantages and disadvantages. We summarize results for models of the LiH, H_2O, and N_2 molecules accounted in Reference 1.

In Figure 2.1, the convergence of the iQCC procedure utilizing various operator pools is compared for LiH in the STO-3G basis set, utilizing a CAS(2e, 3o) active space, and for H_2O in the 6-31G basis set using a CAS(4e, 4o) active space. For H_2O, a bond angle of 107.6°

[‖] Given two Pauli terms $\hat{P}_i$ and $\hat{P}_j$ uniformly sampled from $\{\hat{P}_\alpha\}_{\alpha=1}^{4^N-1}$, they commute and anticommute with equal probabilities 0.5, where $[\hat{P}_i, \hat{P}_j]$ gives Pauli term $2\hat{P}_i\hat{P}_j$ in the case they anticommute. Hence,

$$[\hat{H}, \hat{P}_\alpha] = \sum_{i=1}^{M} c_i [\hat{P}_i, \hat{P}_\alpha] \tag{2.26}$$

gives rise to $M/2$ Pauli terms in expectation, for uniformly sampled $\hat{P}_\alpha$ and $\{\hat{P}_i\}_{i=1}^{M}$, barring any linear dependencies in the anticommutative $\hat{P}_i\hat{P}_\alpha$ instances.

[**] The attenuation can be attributed to 1) terms entering the newly dressed Hamiltonian have coefficient weighted by $\sin(\tau_\alpha^*)$ by inspection of Eq.(2.24), and 2) generally, the optimal values of τ_α^* are decreasing as the number of iterations increases. Consideration 2 can be understood through the fact that individual $\hat{P}_\alpha$ are selected for based on an energy lowering criterion, and hence typically the initial $\hat{P}_\alpha$ will be responsible for generating the most energetically important electronic configurations when acting on $|\phi_0\rangle$. Such electronic configurations often possess high weight in the exact FCI expansion, leading to high optimal amplitudes τ_α^*. Once such strongly correlated configurations have been generated, further iQCC iterations begin to account moreso for the dynamical correlation, where τ_α^* may be near zero.

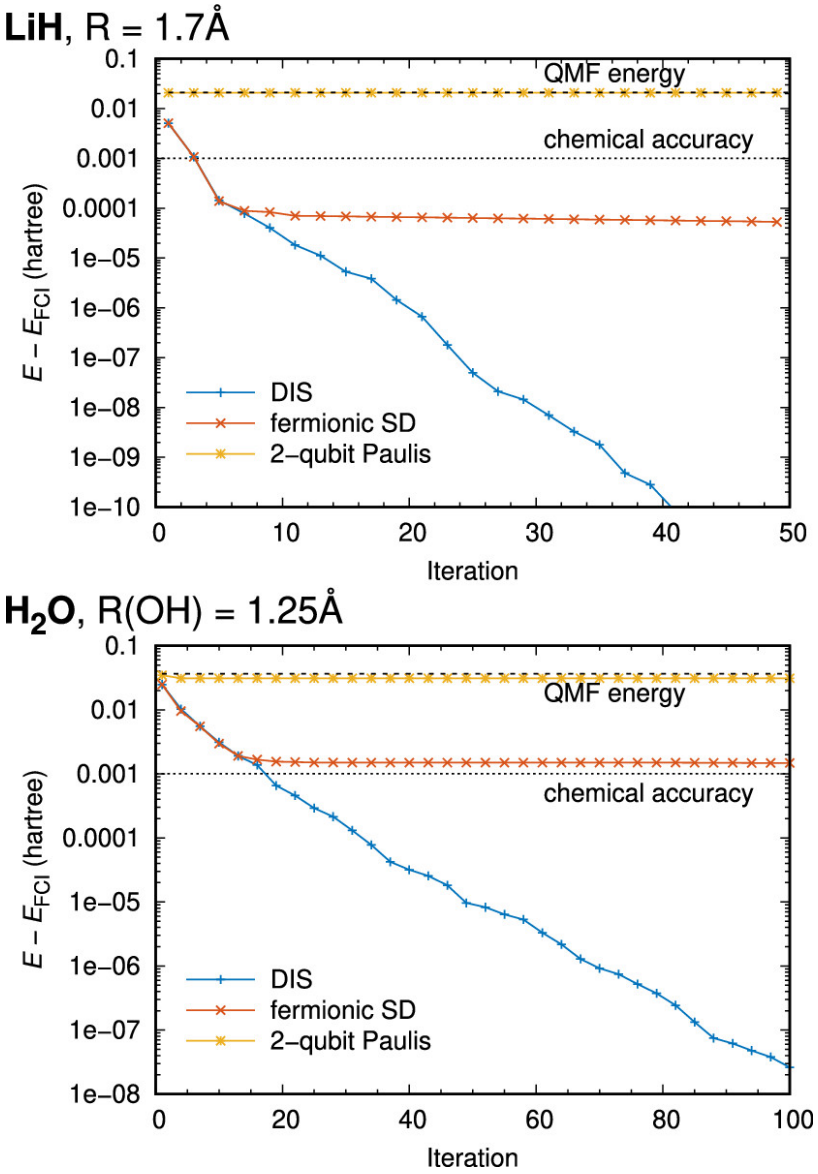

Figure 2.1: Energy convergence of a single-generator ($N_g = 1$) per iteration iQCC procedure for various choices of generator pools: "fermionic SD" is the pool of Pauli terms that arise from fermion-to-qubit mapping of single and double excitation operators. The "2-qubit Pauli" pool is the set of all two-qubit Pauli terms. The direct interaction space (DIS) pool is sampled from by constructing a Pauli term from $\vec{\mu}^{(i)}$ with highest g_i in Eq.(2.22).

was used. Specifically, the iQCC procedure utilizing the DIS construction to identify the top gradient $\mathfrak{su}(2^{N_q})$ basis generators is compared against two restricted operator pools. The "fermionic SD" pool consists of all Pauli terms arising from the fermion-to-qubit mapped single and double excitation/de-excitations, see Eqs. (1.33) and (1.34). The "2-qubit Paulis" pool consists of all possible Pauli terms with support 2, e.g., $\hat{x}_p\hat{y}_q$, $\hat{y}_p\hat{z}_q$, for $p \neq q$. These results include a QMF parameterization of the reference state, and hence the reference includes full flexibility in terms of single-qubit rotations. At each iteration, a single Pauli

term ($N_g = 1$) is selected with highest gradient value from the respective pool. Notably, sampling from either the fermionic SD pool or 2-qubit Pauli pool eventually leads to local energy minima. In the case of sampling from the 2-qubit Pauli pool, this local minima is only a small improvement relative the optimized QMF energy for both examples. In contrast, by utilizing the DIS to freely sample a basis generator of $\mathfrak{su}(2^{N_q})$ of highest possible gradient, the trial energy is systematically lowered far beyond chemical precision.

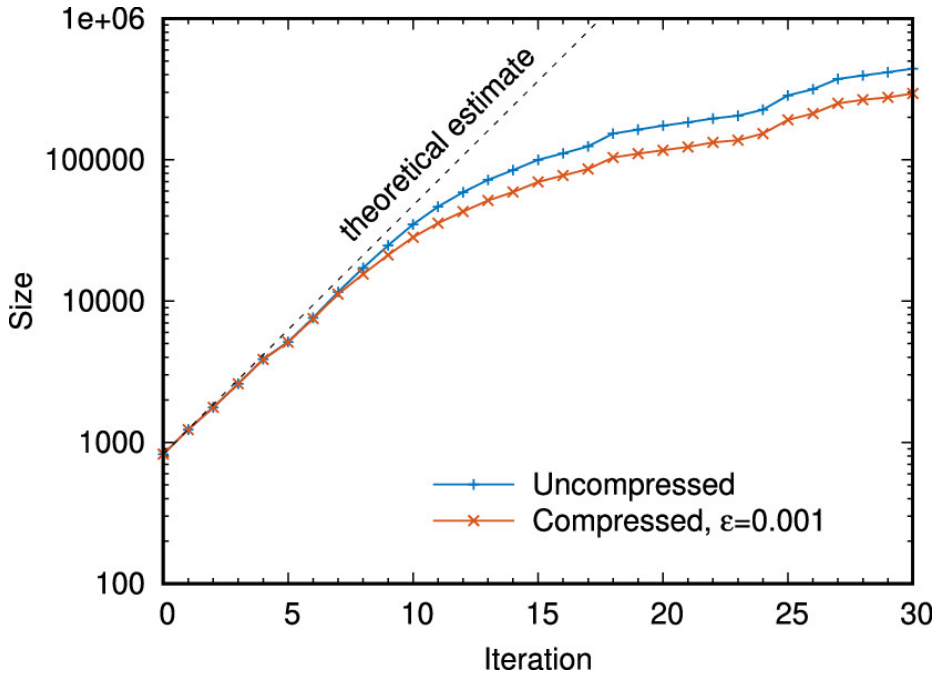

Figure 2.2: Scaling of the number of Hamiltonian terms ("Size") per iQCC iteration, utilizing no term compression, and with compression $\epsilon = 10^{-3}$ Hartree. For the $N_g = 1$ employed, the iteration K is equal to the total number of generators used, hence the theoretical estimate line equals $M(3/2)^K$. The initial qubit-mapped electronic Hamiltonian has $M = 825$ terms.[1]

An analysis of iQCC Hamiltonian growth is given in Figure 2.2 for the N_2 molecule in the cc-pVDZ basis set utilizing a CAS(10e, 8o) basis set. Again, a single generator, $N_g = 1$, is used per iteration (K). The Hamiltonian term counts utilizing no numerical compression are compared with using a compression of $\epsilon = 10^{-3}$, where iQCC Hamiltonian terms of coefficient magnitude less than ϵ are pruned. The theoretical estimate of $M(3/2)^K$ gives reasonable estimates for Hamiltonian term counts for the first 10 iterations, at which point the proliferation of new terms begins to decelerate. By $K = 30$, utilizing compression $\epsilon = 10^{-3}$ yields effective Hamiltonians of ~ 1.5 times fewer terms, with negligible impact on trial energies.

2.6 Conclusions

Numerical experiments indicate that the iQCC procedure converges to the exact energy utilizing a fixed-size ansatz per iteration when the DIS is used as a pool of generators. The robust convergence when utilizing the DIS pool is intuitive: we constructively obtain the highest possible gradient generators from a complete operator pool, the basis generators of $\mathfrak{su}(2^N)$. Hence, any elements of a smaller subset of Pauli terms (such as those Pauli terms arising in the fermionic SD pool) can be identified to be in the DIS. However, the DIS makes no restraint in terms of Pauli term weight, or the rank of excitation from the Pauli term's fermionic origin. The robust energetic convergence includes the case of using only a single $N_g = 1$ generator included at a time, avoiding the necessity of quantum circuit optimization altogether. Since the optimization of the iQCC energy is the only step which may benefit from utilization of a quantum device, e.g., when $N_g > 1$, this leads to a quantum-inspired fully classical parameter regime for the iQCC algorithm. Compared to other competitive adaptive approaches, such as the ADAPT [45] and qADAPT [48] methods, the iQCC approach avoids the growing of a wavefunction ansatz requiring a potentially lengthy circuit description on a quantum device.

However, such advantages come at the cost of a rather steep increase in Hamiltonian terms. The scaling of an iQCC step, implemented as a near-term quantum algorithm, as well as a fully classical algorithm, both depend on the size of the current step iQCC Hamiltonian. While numerical evidence supports that the formal exponential scaling of the iQCC Hamiltonians is pessimistically large beyond the initial iterations in practice, the effective iQCC Hamiltonians still become considerably larger than their initial form. The rapid onset of terms in the initial dressings is considered to be the major computational bottleneck of the iQCC procedure as presented. In Chapters 3 and 4 we present exact algebraic approaches to mitigate the growth of effective iQCC Hamiltonians, as opposed to only approximate numerical compression techniques.

Chapter 3

iQCC with involutory linear combinations of Pauli terms

3.1 Overview

In the preceding chapter, we highlighted that the primary computational challenge in the iterative qubit coupled cluster (iQCC) procedure arises due to the rapidly escalating number of terms. Here, we present an extension of the iQCC procedure designed to mitigate the scaling of the number of Hamiltonian terms. The fundamental concept involves the combining of multiple Pauli terms into a single unitary generator, while maintaining the involutory property. By combining N_g Pauli terms into an involutory linear combination (ILC), we can achieve a significant reduction in the scaling of Hamiltonian terms, resulting in a more favorable complexity of $O(N_g^2)$ instead of the original $O((3/2)^{N_g})$ scaling for the canonical iQCC procedure. This approach is referred to as iQCC with involutory linear combinations (iQCC-ILC), and its detailed explanation follows in the subsequent sections of this chapter. The iQCC-ILC method is numerically compared to the standard iQCC approach, and is seen to give approximately the same quality in energetic errors, while yielding significantly smaller effective Hamiltonians for sufficiently large N_g.

3.2 The iQCC-ILC formalism

Motivated by the unitarity and finite BCH expansion of the QCC dressing, along with its success in lowering circuit depth requirements,[1] we aim to reduce the number of terms introduced in the Hamiltonian when dressed by a unitary similar to $\hat{U}_{\mathrm{QCC}}$ by constructing linear combinations of Pauli terms that are involutory. Using the involutory property of Pauli terms, one can construct an *involutory linear combination* (ILC) from set of Pauli

terms $\mathcal{A} = \{\hat{P}_\alpha\}_{\alpha=1}^{N_g}$,

$$\left(\sum_{\hat{P}_\alpha \in \mathcal{A}} d_\alpha \hat{P}_\alpha\right)^2 = \hat{1}, \tag{3.1}$$

under conditions of normalization,

$$\sum_{\alpha=1}^{N_g} d_\alpha^2 = 1, \tag{3.2}$$

and that all $\hat{P}_\alpha \in \mathcal{A}$ are mutually anti-commutative,

$$\{\hat{P}_\alpha, \hat{P}_\beta\} = 0 \ \forall \ \hat{P}_\alpha, \hat{P}_\beta \in \mathcal{A}. \tag{3.3}$$

With the two conditions satisfied, Eq. (3.1) allows for the encoding of a unitary linear expansion of identity and N_g mutually anti-commuting Pauli terms as the exponent

$$\hat{U}_{\text{ILC}} = \exp\left(-i\tau \sum_{\hat{P}_\alpha \in \mathcal{A}} d_\alpha \hat{P}_\alpha\right) \tag{3.4}$$

$$= \cos(\tau)\hat{1} - i\sin(\tau) \sum_{\hat{P}_\alpha \in \mathcal{A}} d_\alpha \hat{P}_\alpha. \tag{3.5}$$

Satisfying Eq. (3.3) such that the included $\hat{P}_\alpha$'s are variationally significant can be accomplished using the freedom in choosing the $\hat{Z}_\alpha$ part for a $\hat{X}_\alpha$ of high g_α in the DIS (recall Eq.(2.22)). Namely, for every X_α in the DIS, there exists a set of 2^{N_q-1} operators of identical gradient magnitude, which can be generated by all possible choices of $\hat{Z}_\alpha$, such that $\hat{X}_\alpha$ and $\hat{Z}_\alpha$ have odd support. This vast freedom in choosing $\hat{Z}_\alpha$ parts for $\hat{X}_\alpha$'s corresponding to large g_α allows us to find mutually anti-commuting $\hat{P}_\alpha$'s of high gradient magnitude. It is hence possible to construct the desired ILC, up to some restrictions. The ILCs may not be made arbitrarily large, Sankar and van den Berg provided a rigorous upper bound of $2N_q+1$ for the cardinality of sets of mutually anti-commuting N_q-qubit Pauli terms. [66] Further, because generators from the DIS are additionally subjected to a condition that they must have non-zero energy gradients and hence include an odd number of $\hat{y}_i$ operators, the general estimate on maximal ILCs from the DIS is lowered to $2N_q - 1$. Our strategy is to pick $N_g \leq 2N_q - 1$ top-gradient $\hat{X}_\alpha$'s of the DIS, and attempt to construct an anti-commutative set by choice of their $\hat{Z}_\alpha$ parts. In Appendix B we present a procedure for generating sets of anti-commutative elements from the desired set of $\hat{x}$-strings, it requires not more than $O(N_q^6)$ binary arithmetic operations on a classical computer.

Notably, all Pauli terms characterized by the same $\hat{X}_\alpha$ part are mutually commutative under the restraint of possessing an odd number of $\hat{y}$ instances in $\hat{X}_\alpha\hat{Z}_\alpha$. Hence, mutually

anti-commutative Pauli terms of non-zero energy gradient g_α are necessarily possessing different $\hat{X}_\alpha$'s. In obtaining the optimized iQCC-ILC transformation, Eq. (3.4) is optimized with respect to free parameter τ and constrained parameters $\{d_\alpha\}_{\alpha=1}^{|\mathcal{A}|}$ subject to Eq. (3.2). Due to the linearity of the ILC ansatz, Eq. (3.5), one can efficiently obtain the optimal parameters using the linear variational method on a classical computer (further details are in Appendix C).

3.2.1 The iQCC-ILC transformation

The linearity of Eq. (3.5) has the main advantage of encoding N_g high-gradient Pauli terms, generating energetically important configurations as determined by the iQCC screening, while its expansion contains $N_g + 1$ terms as opposed to $O(2^{N_g})$ for the standard QCC unitary 2.1. A transformation of the Hamiltonian with a unitary of form Eq. (3.4) gives

$$\begin{aligned}\hat{U}_{\mathrm{ILC}}\hat{H}\hat{U}_{\mathrm{ILC}}^{\dagger} = &\cos^2(\tau)\hat{H} - \frac{i}{2}\sin(2\tau)\sum_{\alpha=1}^{N_g} d_\alpha[\hat{H},\hat{P}_\alpha] \\ &+ \sin^2(\tau)\sum_{\alpha=1}^{N_g}\sum_{\beta=1}^{N_g} d_\alpha d_\beta \hat{P}_\alpha\hat{H}\hat{P}_\beta. \end{aligned} \tag{3.6}$$

At worst, each commutator $[\hat{H},\hat{P}_\alpha]$ contains the same number of terms as $\hat{H}$, and therefore the summation over N_g commutators in Eq. (3.6) contribute a multiplicative growth factor of N_g at most. By multiplicative growth factor of N_g, it is meant that the resulting $\hat{U}_{\mathrm{ILC}}\hat{H}\hat{U}_{\mathrm{ILC}}^{\dagger}$ will contain $N_g M$ terms for $\hat{H}$ possessing M Pauli terms. Rewriting the double summation in Eq. (3.6) as

$$\begin{aligned}\sum_{\alpha=1}^{N_g}\sum_{\beta=1}^{N_g} d_\alpha d_\beta \hat{P}_\alpha\hat{H}\hat{P}_\beta = &\sum_{\alpha}^{N_g} d_\alpha^2 \hat{P}_\alpha\hat{H}\hat{P}_\alpha \\ &+ \sum_{\alpha>\beta}^{N_g} d_\alpha d_\beta \left(\hat{P}_\alpha\hat{H}\hat{P}_\beta + \hat{P}_\beta\hat{H}\hat{P}_\alpha\right), \end{aligned} \tag{3.7}$$

it is clear that the diagonal summation does not introduce any new terms in $\hat{H}$. This is a consequence of the commutativity/anti-commutativity of all Pauli terms and the involutory property of $\hat{P}_\alpha$'s. The $\hat{P}_\alpha\hat{H}\hat{P}_\beta$ and $\hat{P}_\beta\hat{H}\hat{P}_\alpha$ can be seen to generate the same Pauli terms up to sign differences, and hence the last summation in Eq. (3.7) introduces a $N(N-1)/2$ multiplicative growth factor at worst. The total worst-case multiplicative growth factor G_{worst} for dressing with iQCC-ILC unitary of N_g Pauli terms is

$$G_{\mathrm{worst}}(N_g) = 1 + N_g + N_g(N_g-1)/2 = \frac{1}{2}\left(N_g^2 + N_g + 2\right). \tag{3.8}$$

The average case scaling may be approximately obtained under the assumption that all Pauli products in the Hamiltonian, along with the Pauli terms entering $\hat{U}_{\mathrm{ILC}}$, are uniformly sampled from the set of 4^{N_q} Pauli products. An N_q-qubit Pauli product commutes with half of all 4^{N_q} Pauli products while anti-commuting with the remaining half, leading to the summation over $[\hat{H}, \hat{P}_\alpha]$ in Eq. (3.6) contributing an $N_g/2$ growth factor in expectation. By similar reasoning, each $\hat{P}_\alpha \hat{H} \hat{P}_\beta + \hat{P}_\beta \hat{H} \hat{P}_\alpha$ in the second summation in Eq. (3.7) contributes a growth factor of $1/2$ on average if $[\hat{P}_\alpha, \hat{P}_k] = 0$ and $[\hat{P}_\beta, \hat{P}_k] = 0$ for a given $\hat{P}_k$ in $\hat{H}$ with independent probabilities $1/2$, leading to the summation Eq. (3.7) contributing a growth factor of $N_g(N_g - 1_g)/4$. The total average-case multiplicative growth factor G_{avg} is then

$$G_{\mathrm{avg}}(N_g) = \frac{1}{4}\left(N_g^2 + N_g + 4\right) = \frac{G_{\mathrm{worst}}(N_g) + 1}{2}. \tag{3.9}$$

While the Pauli products in $\hat{H}$ along with selected $\hat{P}_\alpha$ admit more structure than being uniformly random in actuality, we numerically find that Eq. (3.9) yields excellent Hamiltonian growth estimates in Sec. 3.3.

3.2.2 The iQCC-ILC procedure

We describe a high-level overview of the complete iQCC-ILC method herein. The procedure for obtaining the optimized iQCC-ILC transformation, given an input fermion-to-qubit mapped electronic Hamiltonian $\hat{H}$, and number of Pauli terms N_g to enter $\hat{U}_{\mathrm{ILC}}$, is outlined herein.

1. Obtain the DIS $\mathcal{D}_H$ by the generator screening procedure [see Section 2.2].
2. Generate mutually anti-commuting subset $\mathcal{A}$ from $\mathcal{D}_H$ with $|\mathcal{A}| = N_g \leq 2N_q - 1$ using techniques outlined in Appendix B.
3. Obtain the optimal N_g free parameters for $\hat{U}_{\mathrm{ILC}} = \exp\Big(-i\tau \sum_{\hat{P}_\alpha \in \mathcal{A}} d_\alpha \hat{P}_\alpha\Big)$ for qubit mean-field reference $|\phi_0\rangle$ with optional qubit mean-field relaxation using techniques outlined in Appendix C.
4. Dress the Hamiltonian as $\hat{H}' = \hat{U}^\dagger_{\mathrm{ILC}}(\tau, \boldsymbol{d}) \hat{H} \hat{U}_{\mathrm{ILC}}(\tau, \boldsymbol{d})$ [Eq. (3.6)] using the fixed ILC parameters found in the previous step.

The resulting Hamiltonian $\hat{H}'$ can then be used for an ensuing VQE treatment. Optionally, the iQCC-ILC dressing procedure can be repeated, taking the resulting $\hat{H}'$ of Step 4 as an updated input to Step 1. While $\hat{H}'$ is only $O(N_g^2)$ times larger than $\hat{H}$ in the number of its Pauli terms, repeating the iQCC-ILC dressing procedure K times will inevitably result in a final transformed Hamiltonian $O(N_g^{2K})$ times larger than the initial $\hat{H}$. Hence the number of terms in the final transformed Hamiltonian still scales exponentially in the number of iterations K. However, numerical simulations suggest a large portion of electronic correlation can be captured for reasonably low values of K, as presented in Section 3.3. Prior

to running the iQCC-ILC procedure, one can get an accurate estimate of the number of Pauli terms in $\hat{H}'$ by multiplying the number of terms in the initial Hamiltonian by G_{avg} [Eq. (3.9)]. If one wishes to employ exponential procedure of K repeated transformations, the growth factor of the final Hamiltonian can be approximately taken as G_{avg}^K. Alternatively, K may be determined *a posteriori* by convergence of the optimized iQCC-ILC energies [see Eq. (C.7)] or gradient magnitudes present in $\mathcal{D}_H$. For instance, if the iQCC-ILC energy difference between consecutive iterations falls below a given threshold, or the maximum gradient magnitude for $\mathcal{D}_H$ in Step 1 is below a threshold, then exit the dressing procedure.

Due to the iQCC-ILC transformation's linear parameterization, it does not depend on the order of $\hat{P}_\alpha$'s. However, if one is to re-calculate the DIS along a potential energy surface (PES), it is possible that the top N partitions change, resulting in a different set of Pauli terms selected as unitary generators at adjacent nuclear coordinates along the PES. Such a problem occurs with the standard iQCC procedure, and can potentially lead to significant kinks in the PES if convergence is poorly achieved, an inherent disadvantage of adaptive ansätze. To avoid the possibility of energetic kinks arising from ansatz adaptation along the PES, one can perform DIS calculation and subsequent ILC construction at a single geometry, and use the resulting ansatz for optimization over the entire PES. The point in which the DIS is computed can be taken to be any nuclear geometry where electron correlation is expected to be significant. Since the gradient magnitudes are perturbative measures of correlation energy, one can calculate the DIS magnitudes along the PES, and perform ILC construction using the highest N_g partitions where gradient magnitudes are maximized.

3.3 Results

Within this section, we numerically elucidate various properties of the introduced iQCC-ILC algorithm for molecules in variously sized active spaces. The capability of the iQCC-ILC transformation employed as a classical overhead to subsequent VQE application is illustrated for the potential energy curves (PESs) of LiH dissociation and symmetric bond stretch of H_2O in the STO-3G and 6-31G bases respectively. The increase in the number of Hamiltonian Pauli terms resulting from random iQCC-ILC transformations is benchmarked against randomized QCC transformations of similar specification for N_2 in the cc-pVDZ basis. Finally, a Hamiltonian growth comparison is conducted using optimally selected generators for the iQCC-ILC and standard QCC ansätzes, using their optimized amplitudes for dressing, along with a benchmark of their corresponding energy estimates, for H_2O in the 6-31G(d) basis. When Hamiltonian Pauli term counts are reported, only terms with coefficient magnitudes exceeding 10^{-8} are accounted for, i.e., a compression of $\epsilon_c = 10^{-8}$ is used.

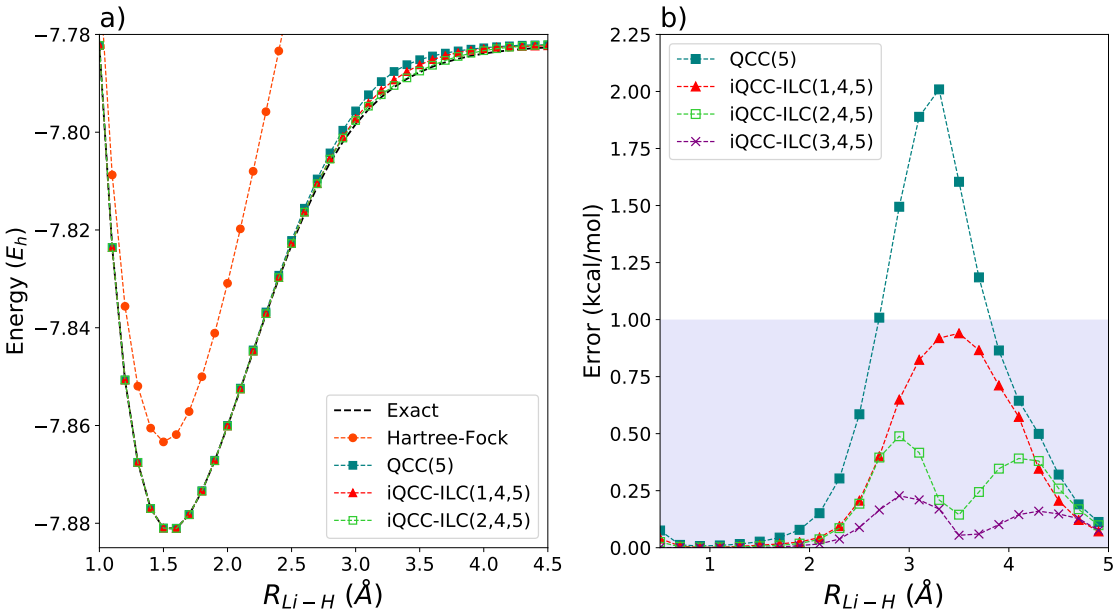

Figure 3.1: a) PESs for the ground state CAS(2e, 2o) LiH dissociation in different methods. QCC(5) ansatz uses the bare qubit Hamiltonian and 5 Pauli terms ($\hat{x}_1\hat{y}_3$, $\hat{x}_1\hat{y}_3\hat{x}_4$, $\hat{x}_1\hat{x}_2\hat{y}_3x_4$, $\hat{x}_1\hat{y}_2\hat{x}_3$, and $\hat{x}_1\hat{y}_2\hat{x}_3\hat{x}_4$). b) Ground state energy errors along the LiH dissociation for different methods. The shaded area denotes errors within chemical accuracy.

To assess the capabilities of the iQCC-ILC transformation in increasing the electronic energy estimates for VQE optimizations using limited quantum resources, we perform the iQCC-ILC procedure as an overhead to standard QCC VQE application. We use iQCC-ILC transformed Hamiltonian in obtaining simulated QCC VQE energies along the PESs for dissociation of LiH and symmetric dissociation of H_2O in the STO-3G and 6-31G basis sets respectively. Active space fermionic Hamiltonians for the LiH and H_2O systems were generated and transformed to their qubit space representations by parity and BK transformations respectively. An active space of 2 electrons in 3 molecular orbitals, CAS(2e, 3o) was employed for LiH, and CAS(4e, 4o) for H_2O. Following qubit tapering procedures, [67] the resulting LiH and H_2O Hamiltonians act on 4- and 6-qubit spaces respectively.

While repeating the iQCC-ILC procedure is bound to result in an exponential blowup in the number of Pauli terms of the dressed Hamiltonian, it is demonstrated that it is possible to increasingly improve VQE energy estimates limited by fixed-depth quantum circuits. We see that generally the first iQCC-ILC dressing will capture the majority of electronic correlation, and hence the polynomially scaling $K = 1$ procedure is of the greatest practicality. We denote the sequence of K iQCC-ILC dressings of $N_g^{(d)}$ Pauli terms each followed by subsequent VQE circuit simulation using QCC ansatz of $N_g^{(c)}$ Pauli terms as iQCC-ILC(K, $N_g^{(d)}$, $N_g^{(c)}$). [46]

With initial LiH qubit Hamiltonians of 100 terms, a single dressing produces Hamilto-

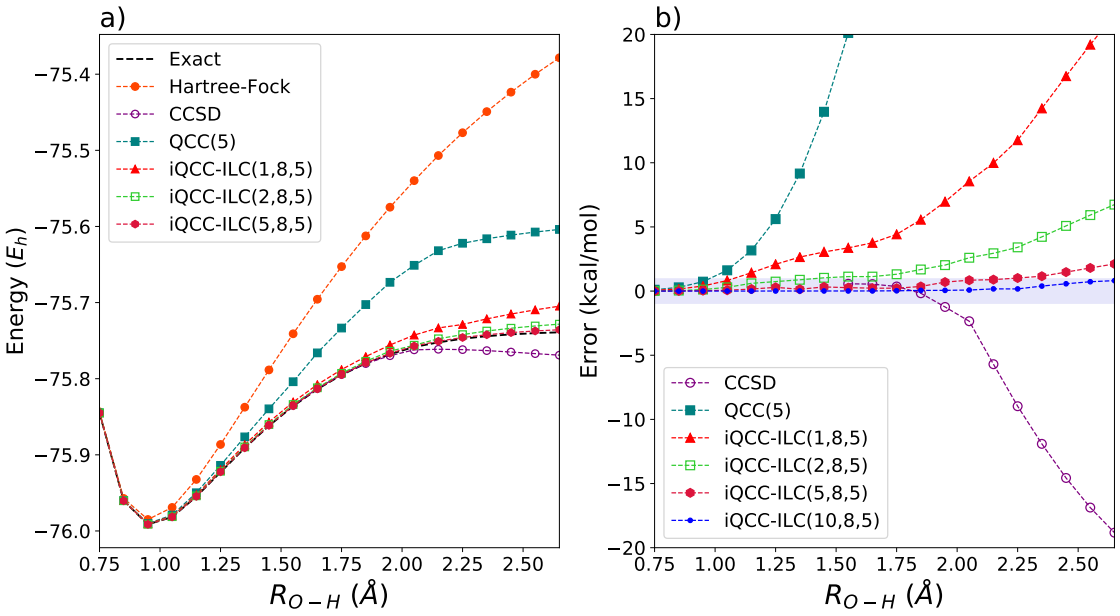

Figure 3.2: a) PESs for the ground state of CAS(4e, 4o) H_2O along the symmetric bond stretch in different methods. QCC(5) employs the spin-penalized Hamiltonian and 5 Pauli terms ($\hat{x}_2\hat{y}_3\hat{x}_5\hat{x}_6$, $\hat{x}_2\hat{y}_3\hat{x}_4\hat{x}_5$, $\hat{x}_1\hat{y}_3$, $\hat{x}_1\hat{y}_2\hat{x}_5\hat{x}_6$, and $\hat{x}_1\hat{y}_2\hat{x}_4\hat{x}_5$). b) Ground state energy errors along the symmetric bond stretch for different methods. The shaded area denotes errors within chemical accuracy.

nians consisting of 135 terms. A second and further dressings yield a transformed Hamiltonian of 136 terms which appears to be the limiting value. This is the result of conserving the time-reversal symmetry of the original Hamiltonian by the iQCC-ILC transformations. Time-reversal symmetry requires even number of $\hat{y}$ operators in individual Pauli products and reality of their coefficients. The number of possible even $\hat{y}$-parity N_q-qubit Pauli terms is

$$\sum_{i=0}^{\left\lfloor \frac{N_q}{2} \right\rfloor} \binom{N_q}{2i} 3^{N_q-2i}, \tag{3.10}$$

which gives 136 for the $N_q = 4$ case.

At the highly extended geometries of the H_2O model, a triplet solution approaches the ground state singlet, which can lead the iQCC procedure to the higher spin solution. To improve convergence to the singlet ground state for H_2O, all iQCC/iQCC-ILC calculations were performed starting with a spin-penalized Hamiltonian,

$$\hat{H}_s = \hat{H} + \frac{\mu}{2}\hat{S}^2, \tag{3.11}$$

where μ is a positive constant ($\mu = 0.5E_h$). This extra term serves as a penalty on non-

singlet spin contamination during optimization. Since $\hat{S}^2$ is a two-body operator, $\hat{H}_s$ and $\hat{H}$ contain roughly the same number of terms.

The initial Hamiltonian $\hat{H}_s$ for water contains 165 terms. The transformed Hamiltonian dressed with a single iQCC-ILC unitary of 8 Pauli terms obtained in the iQCC-ILC(1, 8, 5) procedure contains 995 terms, demonstrating a ~ 6 times growth factor. This is quite low in relation to the average growth factor $G_{\text{avg}} = 19$ predicted by Eq. (3.8) due to near-saturation of maximum Hamiltonian terms. Systematically improved energy estimates are obtained with the iQCC-ILC(K, 8, 5) for increasing K. The iQCC-ILC procedure with $K = 10$ transformations demonstrates microHartree accuracy for the majority of the PES, with chemical accuracy achieved for the entire considered reaction coordinate. The iQCC-ILC(2, 8, 5) procedure is already seen to energetically outperform the coupled cluster with single and double excitations (CCSD) result at extended geometry (Figure 3.2b), the latter produces significant non-variational behaviour.

The iQCC and iQCC-ILC transformations with the same $N_g^{(d)}$ and K applied to these systems give comparable Hamiltonian growth. The comparable scaling of the Hamiltonian growth is presumed to be a symptom of the low upper bound on number of terms which can enter the transformed Hamiltonian for $N_q \leq 6$. To numerically assess the Hamiltonian growth resulting from ILC unitary transformations compared with standard QCC product transformations, we employ transformations on an N_2 electronic Hamiltonian in the cc-pVDZ basis in an active space of 6 electrons in 6 spatial orbitals, CAS(6e,6o). Molecular spin orbitals are enumerated using the same-spin grouping, and the 12 qubit Hamiltonian consisting of 247 Pauli terms is obtained via parity transformation of the canonical fermionic creation/annihilation operators. The typical growth incurred by the dressings is numerically assessed by performing uniformly random QCC and ILC transformations on the initial Hamiltonian with varying number of Pauli terms (N) entering the QCC/ILC unitaries, and plotted in Figure 3.3. A random QCC transformation of N Pauli terms is obtained by uniformly choosing N Pauli terms of odd y-parity, and uniformly setting their associated amplitudes in the interval $[0, 2\pi)$. A random ILC transformation of N anti-commuting Pauli terms is obtained from the following procedure: 1) N uniformly random bit-strings of length N_q are generated and used to represent a set of N DIS flip indices (describing placement of $\hat{x}_i$'s and $\hat{y}_i$'s), 2) the procedure of Appendix B is then used to find a set of N anti-commuting Pauli terms from the flip-indices by solving Eq. (B.4), 3) finally the global amplitude of the ILC unitary is then selected uniformly at random, $\tau \in [0, 2\pi)$, and the remaining parameters $\{d\}_{i=1}^{N}$ are taken to be the components of a random N-dimensional unit vector. The averages and standard deviations in Figure 3.3 are obtained from 10 trials of the uniformly random dressing procedures for each value of N. The standard deviations for both the QCC and ILC dressings are remarkably low relative to the averages with only 10 trials performed, suggesting the number of Pauli terms in the dressed Hamiltonians is predominantly determined by the number of Pauli terms entering the unitary, rather

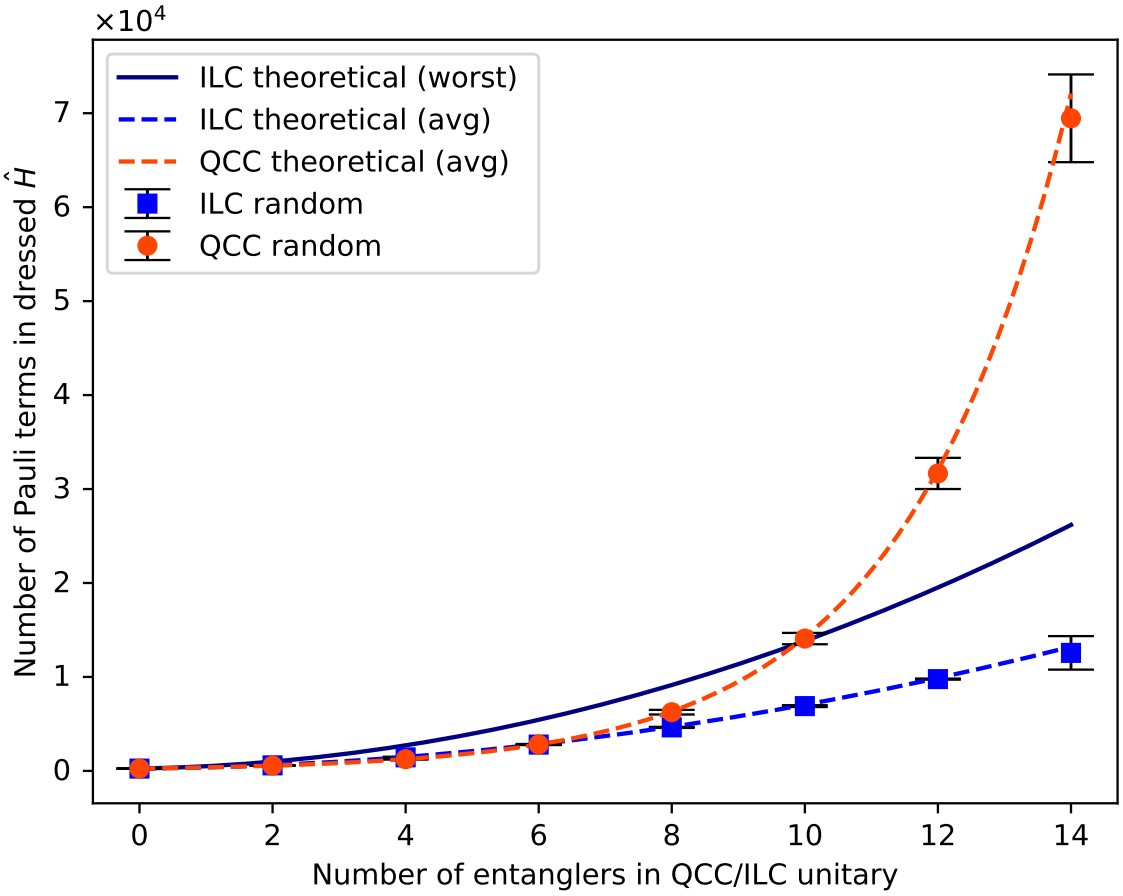

Figure 3.3: Comparison of theoretical and random case Pauli term growth of the CAS(6e, 6o) N_2 initial Hamiltonian ($\hat{H}_{N_2}$) in the cc-pVDZ basis after QCC(N_g) and ILC(N_g) unitary transformations. Points denote the average number of Pauli terms in the dressed Hamiltonians, and error bars denote standard deviations. The number of Pauli terms in $\hat{H}_{N_2}$ is $M = 247$. The worst and average case theoretical estimates for the ILC transformation are given by $MG_{\text{worst}}(N_g)$ and $MG_{\text{avg}}(N_g)$ (see Eqs. (3.8) and (3.9)) respectively. The average case theoretical QCC transformation size estimate is $M(3/2)^{N_g}$.

than which specific Pauli terms are chosen and the values of their associated amplitudes. Therefore, the average-case multiplicative growth factor Eq. (3.9) can be used to obtain estimates for the size of iQCC-ILC transformed Hamiltonians to high accuracy.

To assess how the ILC energy compares to QCC energies in regimes where the ILC dressed Hamiltonians become sufficiently smaller than QCC dressed Hamiltonians, we perform benchmarking using a full valence active space H_2O electronic Hamiltonian in the 6-31G(d) basis. From the total 19 atomic orbitals, only the 12 *sp*-type HF molecular orbitals were included to form the active space of CAS(8e, 12o), i.e. freezing the oxygen $1s$ core and the remaining d-character virtual orbitals. The H−O−H angle was set to 107.6°, and a symmetric O−H bond length of $R = 1.5$ Å was used. The resulting 24 qubit Hamiltonian consisted of 8921 Pauli terms after the JW mapping.

We compare optimized ILC energies and transformed Hamiltonian growth factors with those of various optimal QCC ansätze in Figure 3.4. An optimal QCC ansatz of N_g Pauli terms, QCC(N_g), is formed by selecting one element out of $O(2^{N_q-1})$ choices from each of the N_g DIS partitions of largest gradient magnitude, and ordered by ascending gradient magnitude. While any choice of Pauli terms from the top N DIS partitions is considered to be optimal by the first-order gradient heuristic, the final optimized energy and transformed Hamiltonian size will be a function of the specific Pauli term selections. Thus, in order to assess how the ILC transformation of N_g anti-commuting Pauli terms, ILC(N_g), compares to analogous QCC(N_g) transformations, we construct 500 samples of top-gradient QCC(N_g) ansätze by randomly selecting a DIS element from each of the top N_g partitions. All ansätze use the gradient magnitude ordering. Each resulting QCC energy functional was optimized using the limited-memory Broyden-Fletcher-Goldfarb-Shanno (L-BFGS) algorithm. [29]

The average performance of the first-order optimal QCC(N) ansätze compared with the ILC(N) ansatz are displayed in Table 3.1. Pauli term counts for the dressed Hamiltonians using the ILC(N) transformation become drastically lower than those for the average of first-order optimal QCC(N) dressing for increasing N, with nearly a factor of 2 reduction in growth for $N = 10$. By extrapolation one can expect more drastic reductions for $N > 10$. The tabulated relative energies indicate that using the ILC transformation can significantly reduce Pauli term growth of the dressed Hamiltonian while robustly maintaining marginally lower energies when compared to the average over first-order optimal QCC unitary ansätze of similar specification for large scale calculations.

While the implemented ILC(8) and ILC(10) unitaries are not first-order optimal, they still provide energies slightly lower than the first-order optimal QCC(8) and QCC(10) averages respectively. Since one cannot find a set of anti-commuting Pauli terms from the top 8 (10) DIS partitions here, a brute force search through the DIS partitions is performed until a set is identified which mutually anti-commuting Pauli terms can be found from, i.e. for which a solution exists in the procedure outlined in Appendix B. The identified set of 8 (10) partitions turn out to have very similar gradient magnitudes to the top 8 (10). The

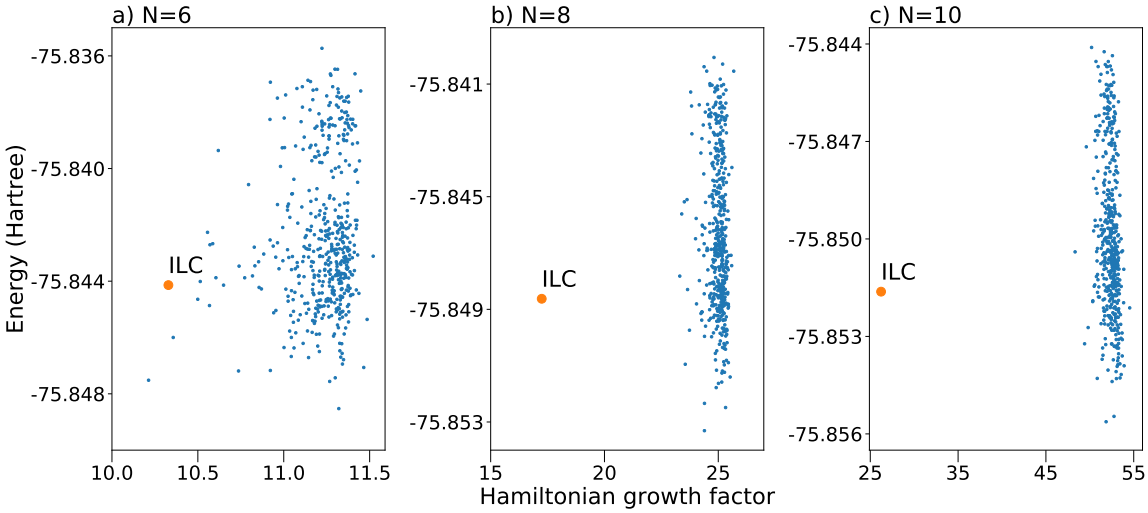

Figure 3.4: Comparison of optimized energies and dressed Hamiltonian growth for the ILC(N_g) ansatz and 500 randomly chosen first-order optimal QCC(N_g) ansätze (blue points) applied to the CAS(8e, 12o) H_2O electronic Hamiltonian in the 6-31G(d) basis at $R = 1.5$ Å. The Hamiltonian growth factors are taken as the ratio of the number of Pauli terms in the dressed Hamiltonian to the number of Pauli terms in the initial Hamiltonian (8921).

similar gradient magnitudes between the set of top N and the set of anti-commutative set-constructable N partitions explains the comparable optimized energies up to a first-order consideration. One can expect this to be a common scenario in the viscinity of orbital quasi-degeneracies: there will be many DIS partitions of quasi-degenerate gradient magnitude, and hence one can find many sets of N partitions that have gradient magnitudes similar to those of the top N set of partitions.

3.4 Conclusions

Within this chapter, we have introduced an alternative unitary ansatz to be used within the iQCC procedure which mitigates Hamiltonian growth. The iQCC-ILC transformation is unitary, allowing for straightforward variational application. Further, the exact transformation using an ILC of N_g Pauli terms results in $O(MN_g^2)$ terms when applied to a qubit-space Hamiltonian of M Pauli terms. Such a quadratic growth factor in N_g is the main advantage over the standard iQCC procedure, which has the unfortunate scaling of $O(M(3/2)^{N_g})$. Furthermore, the linearity of the iQCC-ILC unitary allows for efficient global parameter optimization on a classical computer for general N_g, replacing a potentially burdensome nonlinear multivariable optimization* in the standard iQCC procedure. One can apply the iQCC-ILC optimization procedure as an efficient classical pre-processing technique that alleviates subsequent VQE-style optimization. However, it may also be applied as a purely

*Or a $O(2^{N_g})$ dimensional matrix diagonalization if following the technique described in Section 2.3

Table 3.1: Comparison of the optimized energies and the number of terms in the dressed Hamiltonian for the ILC(N_g) ansatz and the average of the first-order optimal QCC(N_g) ansätze obtained from 500 samples (see Figure 3.4). The energies for the ILC(N_g) ansatz and the average of the QCC(N_g) ansätze are denoted $E_{\mathrm{ILC}(N)}$ and $E^{(avg)}_{\mathrm{QCC}(N)}$. The number of Pauli terms in the dressed Hamiltonians for the ILC(N_g) ansatz and the average of the QCC(N_g) ansätze are denoted $M_{\mathrm{ILC}(N)}$ and $M^{(avg)}_{\mathrm{QCC}(N)}$ respectively.

N_g	$M^{(avg)}_{\mathrm{QCC}(N_g)}$ $\times 10^{-3}$	$M_{\mathrm{ILC}(N_g)}$ $\times 10^{-3}$	$E_{\mathrm{ILC}(N_g)} - E^{(avg)}_{\mathrm{QCC}(N_g)}$ (mHartrees)
6	100	92	-1.72
8	223	154	-2.34
10	467	234	-1.76

classical quantum-inspired approach, and perturbative techniques have since been developed to further increase accuracy in this setting.[68] The iQCC-ILC transformation is seen to aid significantly in yielding accurate energy estimates using fixed circuit depth QCC ansätze particularly in the strongly correlated regime (e.g. symmetric bond stretch in H_2O). In this regime, the reference state $|\phi_0\rangle$ used in the undressed QCC generator selection procedure has small overlap with the highly correlated ground state of interest. This leads to fast growth of the number of necessary Pauli terms in the QCC scheme and to unfavorable circuit depths. The iQCC-ILC transformation was seen to capture significant amounts of correlation energy, particularly in the strongly correlated regime, as seen for the simultaneous bond-dissociation of CAS(4e, 4o) H_2O in the 6-31G basis. Systematic reduction in Hamiltonian Pauli term growth from the iQCC-ILC transformation compared to analogous iQCC dressings was demonstrated via random generator selection and amplitude sampling for CAS(6e, 6o) N_2 in the cc-pVDZ basis, and first-order optimal generator selection with optimized amplitudes for CAS(8e, 12o) H_2O in the 6-31G(d) basis.

Chapter 4

iQCC with growth-mitigating Pauli terms

4.1 Overview

Herein this chapter we present an alternative route to mitigating the onset of new terms in the iQCC effective Hamiltonians. In the iQCC-ILC procedure of Chapter 3, the algorithm seeks to combine multiple high energy gradient $\mathfrak{su}(2^{N_q})$ basis generators into a single involutory generator as a linear combination, facilitating an efficient closed form BCH expansion when dressing the Hamiltonian. While combining multiple Pauli terms into a single involutory generator incurs only a $O(N_g^2)$ growth factor of the Hamiltonian per iteration, as opposed to $O((3/2)^{N_g})$ of the standard iQCC procedure, there is still an exponential dependence on number of iterations K. In the "growth-mitigated" iQCC (GM-iQCC) procedure introduced herein, an alternative heuristic approach to mitigating the proliferation of terms is taken. Due to the vast degree of choices in specifying the $\hat{Z}_\alpha$ part in the $\theta_\alpha \hat{X}_\alpha \hat{Z}_\alpha$ factorized form of $\hat{P}_\alpha$ while keeping the gradient magnitude g_α fixed, the GM-iQCC procedure utilizes a classical algorithm to identify the $\hat{Z}_\alpha$ component which minimizes the onset of new terms upon Hamiltonian dressing, for fixed $\hat{X}_\alpha$, and hence for fixed energy gradient.

4.2 The GM-iQCC formalism

4.2.1 Assessing Hamiltonian growth

Herein we provide a brief derivation of how to quantify the increase in number of terms when dressing a qubit-space Hamiltonian $\hat{H}$ with unitary generated by $\hat{P}_\alpha$. Let $\mathcal{H}^{(K)} = \{\hat{P}_i\}_{i=1}^{M}$ denote the set of Pauli products in $\hat{H}^{(K)}$. By inspection of the iQCC dressing expression Eq.(2.24), new terms in the Hamiltonian transformed with $\exp\left(-i\tau_\alpha^* \hat{P}_\alpha/2\right)$ are introduced only by $[\hat{H}, \hat{P}_\alpha]$. Since two Pauli products either commute or anticommute, $\hat{P}_\alpha \hat{H} \hat{P}_\alpha$ will not introduce Pauli products not already in $\hat{H} = \sum_i \eta_i \hat{P}_i$, since $\hat{P}_\alpha \sum_i \eta_i \hat{P}_i \hat{P}_\alpha = \sum_i \eta_i \theta_i \hat{P}_i$,

where θ_i equals 1 (-1) if $\hat{P}_i$ and $\hat{P}_\alpha$ commute (anticommute). Further, any commutator $[\hat{P}_i, \hat{P}_\alpha]$ in $[\hat{H}, \hat{P}_\alpha] = \sum \eta_i [\hat{P}_i, \hat{P}_\alpha]$ will contribute a new term not found in $\mathcal{H}^{(K)}$ if $\hat{P}_i$ and $\hat{P}_\alpha$ anticommute and $[\hat{P}_i, \hat{P}_\alpha] = 2\hat{P}_i\hat{P}_\alpha$ is not proportional to a term in $\mathcal{H}^{(K)}$. Therefore, we can quantify the growth incurred by transforming the Hamiltonian with $\exp\left(-i\tau_\alpha^* \hat{P}_\alpha/2\right)$ as the number of terms in $[\hat{H}^{(K)}, \hat{P}_\alpha]$ that are not found in $\mathcal{H}^{(K)}$. We denote this quantity as γ_α, which serves as an analytically exact upper bound to the increase in number of terms upon conjugating the current step Hamiltonian with $\exp\left(-i\tau_\alpha^* \hat{P}_\alpha/2\right)$,

$$|\mathcal{H}^{(K+1)}| \leq |\mathcal{H}^{(K)}| + \gamma_\alpha. \tag{4.1}$$

The case of inequality arises when there are term cancellations, or when magnitude of a Pauli product coefficient in the resulting Hamiltonian falls below a precision threshold, and is hence omitted due to numerical compression. Such cases will be dependent on the value of $\tau_\alpha^{'}$, which are not accounted for in the analytical growth consideration.

In the context of an adaptive Hamiltonian transformation, we generally wish to select Pauli products from the pool with high g_α and low γ_α. Similarly to how energetic gradients can not be directly assessed for the exponential number of elements in the Pauli group, assessing growth explicitly for all pool elements is computationally expensive. Instead, we devise efficient classical algorithms for obtaining the pool element with smallest γ_α within a given gradient partition. These algorithms and their derivation are included in Appendix D. For a given partition k, we denote the minimal growth by $\gamma^{(k)}$, i.e., the lowest growth observed in the set of Pauli products with gradient $g^{(k)}$. We can hence use both $g^{(k)}$ and $\gamma^{(k)}$ in determining which partition of the Pauli product pool will be sampled from. In the next section, we introduce a parameterized scoring function which ranks the partitions of the Pauli product pool based on both $g^{(k)}$'s and $\gamma^{(k)}$'s.

4.2.2 Hybrid operator importance measures

The majority of adaptive ansatz schemes, including iQCC, have used proxies for expected energy lowering such as energetic gradients g_α's as the sole importance measure when considering which elements from the operator pool should be included in the ansatz at a given iteration. In the context of an effective Hamiltonian theory, we can introduce another importance measure based on the worst-case increase in the number of Hamiltonian Pauli products upon transformation generated by a given element. To consider both energetic gradients and growth in selecting from the operator pool, we introduce score function $s^{(k)}$ assigned to partitions of the pool. Since energetic convergence will largely be determined by the energetic gradients of the chosen generators, we only consider scoring the P highest gradient pool partitions. We define this cumulative score function as

$$s^{(k)} = a\tilde{g}^{(k)} - (1-a)\,\tilde{\gamma}^{(k)}, \tag{4.2}$$

where $\tilde{g}_k$ and $\tilde{\gamma}^{(k)}$ are normalized energetic gradients and growth, respectively,

$$\tilde{g}^{(k)} = g^{(k)} \left(\frac{\sum_{i=1}^{P} g^{(i)}}{P} \right)^{-1}, \tag{4.3}$$

$$\tilde{\gamma}^{(k)} = \gamma^{(k)} \left(\frac{\sum_{i=1}^{P} \gamma^{(i)}}{P} \right)^{-1}. \tag{4.4}$$

Real parameter $a \in [0, 1]$ in Eq. (4.2) acts as a bias for energetic gradient. For instance, sampling from the highest energy gradient partition regardless of minimum growth estimates is accomplished by setting $a = 1$.

4.2.3 The GM-iQCC procedure

Below, we outline the procedure for *growth mitigated* iQCC, denoted GM(a)-iQCC, where a denotes the gradient bias utilized in scoring function $s(a)$ (see Eq. (4.2)). We use $s(a)$ with several different gradient bias values for assessment. All iQCC calculations herein use $N_{\text{gen}} = 1$, hence the highest gradient partition is always selected from in the canonical scheme. Explicitly, the employed procedure for obtaining the K^{th} step effective Hamiltonian in the GM(a)-iQCC scheme is:

1. Obtain description of DIS for $\hat{H}^{(K-1)}$ (see Section 2.2). Let $\{g^{(k)}\}_{k=1}^{P}$ be the set of the P highest energetic gradient magnitudes.

2. For each of the P highest gradient partitions, obtain the lowest growth incurring element (see Appendix D). In doing so, we obtain the minimal growth value of the top P partitions $\{\gamma^{(k)}\}_{k=1}^{P}$.

3. Compute the scores $\{s^{(k)}\}_{k=1}^{P}$ of the top P partitions using chosen bias a, and computed $\{g^{(k)}\}_{k=1}^{P}$ and $\{\gamma^{(k)}\}_{k=1}^{P}$ (see Eq. (4.2)).

4. Select $\hat{P}_K$ corresponding to the growth minimizing element of the top scoring partition according to $\{s^{(k)}\}_{k=1}^{P}$.

5. Perform optimization

$$\min_{\tau_K} E_K(\tau_K) = \langle 0 | \, e^{i\tau_K \hat{P}_K/2} \hat{H}^{(K-1)} e^{-i\tau_K \hat{P}_K/2} \, |0\rangle \, . \tag{4.5}$$

   Optimized $\min_{\tau_K} E_K(\tau_K)$ is taken as the K^{th} step of the energy estimation.

6. Using $\tau_K^* = \arg\min E_K(\tau_K)$, obtain K^{th} step effective Hamiltonian

$$\hat{H}^{(K)} = e^{i\tau_K \hat{P}_K/2} \hat{H}^{(K-1)} e^{-i\tau_K \hat{P}_K/2}. \tag{4.6}$$

It should be emphasized that Step 4 features the selection of the *growth minimizing element* of the highest scoring partition. For instance, the $s(1)$ parameterization results in the growth minimizing element of the highest energy gradient partition being selected for at each step in the GM(1)-iQCC procedure. This differentiates GM(1)-iQCC from iQCC utilizing the canonical sampling scheme, while also always selects from the highest gradient partition, however with no effort made to minimize growth by the specific Pauli product chosen within the partition.

4.2.4 Obtaining growth-minimizing elements of fixed gradient from the DIS

This section is organized as follows. Firstly, we formulate the conditions for finding Pauli products which do not introduce any new terms in $\hat{H}$ upon similarity transformation with $\exp\Big(-i\tau_\alpha \hat{P}_\alpha/2\Big)$. We refer to such Pauli products as "normalizer" Pauli products. Then, we propose a convenient method of computing the analytical growth of candidate Pauli products, which allows us to find normalizers if they exist for a particular Hamiltonian. If normalizers do not exist, we can use the methods developed herein to identify growth minimizing Pauli products. We then introduce deterministic and probabilistic algorithms for obtaining minimal growth Pauli products with fixed energetic gradient g_α values.

Since new terms arise strictly from commutator $[\hat{H}, \hat{P}_\alpha]$, the condition for $\hat{P}_\alpha$ to be a normalizer Pauli product is

$$[\mathcal{H}, \hat{P}_\alpha] \subseteq \mathcal{H} \cup \{0\}, \tag{4.7}$$

where $\mathcal{H}$ is the set of Pauli products in $\hat{H}$, $\mathcal{H} = \{\hat{P}_i\}_{i=1}^M$. The commutator of $\mathcal{H}$ and $\hat{P}_\alpha$ is defined to produce a set of element-wise commutators, i.e. $[\mathcal{H}, \hat{P}_\alpha] = \{[\hat{P}_1, \hat{P}_\alpha], \ldots, [\hat{P}_M, \hat{P}_\alpha]\}$. Here and throughout this section, we ignore coefficients arising in non-zero commutators between Pauli products. Let $\mathcal{H}_C^{(\alpha)} \subseteq \mathcal{H}$ denote the subset of Hamiltonian Pauli products which commute with $\hat{P}_\alpha$,

$$\mathcal{H}_C^{(\alpha)} = \{\hat{P}_i \in \mathcal{H} \mid [\hat{P}_i, \hat{P}_\alpha] = 0\}, \tag{4.8}$$

and similarly let $\mathcal{H}_A^{(\alpha)} = \mathcal{H}/\mathcal{H}_C^{(\alpha)}$ denote the subset of Hamiltonian terms which anticommute with $\hat{P}_\alpha$. By definition, $[\mathcal{H}_C^{(\alpha)}, \hat{P}_\alpha] = \{0\}$. Commutation of any $\hat{P}_i \in \mathcal{H}_A^{(\alpha)}$ with $\hat{P}_\alpha$ results in $[\hat{P}_i, \hat{P}_\alpha] = 2\hat{P}_i\hat{P}_\alpha$, which anticommutes with $\hat{P}_\alpha$. Therefore condition Eq. (4.7) can be refined as

$$[\mathcal{H}_A^{(\alpha)}, \hat{P}_\alpha] = \mathcal{H}_A^{(\alpha)}. \tag{4.9}$$

Essentially, normalizer $\hat{P}_\alpha$ induces an automorphism $[\cdot, \hat{P}_\alpha]$ on the set of Hamiltonian terms anticommuting with $\hat{P}_\alpha$. Since $[\hat{P}_i, \hat{P}_\alpha] = 2\hat{P}_i\hat{P}_\alpha$ for $\hat{P}_i \in \mathcal{H}_A^{(\alpha)}$, it follows that there exists

$\hat{P}_j \in \mathcal{H}_A^{(\alpha)}$ which is proportional to $\hat{P}_i\hat{P}_\alpha$, if $\hat{P}_\alpha$ is a normalizer Pauli product. Hence, for normalizer $\hat{P}_\alpha$, we must find $|\mathcal{H}_A^{(\alpha)}|/2$ disjoint pairs of terms $(\hat{P}_i, \hat{P}_j)$ such that $\hat{P}_i\hat{P}_j = \hat{P}_\alpha$ up to a multiplicative phase. Since $\hat{P}_j = \hat{P}_i\hat{P}_\alpha$ up to a phase, it also follows that $(\hat{P}_i, \hat{P}_j)$ are an anticommutative pair, and $\hat{P}_\alpha = \hat{P}_i\hat{P}_j$ anticommutes with both $\hat{P}_i$ and $\hat{P}_j$. These facts allow us to derive a classical algorithm to efficiently find growth minimizing terms, as described below.

The principal object used in formulating the algorithm is the multiset of non-zero commutators which can be evaluated between unique pairs formed from set $\mathcal{H}$. Let

$$\mathcal{C} = \left\{ [\hat{P}_i, \hat{P}_j] \mid \hat{P}_i, \hat{P}_j \in \mathcal{H}, i < j \right\} / \{0\}, \tag{4.10}$$

where we again ignore any prefactors arising in nonzero $[\hat{P}_i, \hat{P}_j]$. The requirement of $i < j$ is used to ignore redundant pairs. Once $\mathcal{C}$ is evaluated, we can then check for the multiplicity of each unique $\hat{P}_\alpha \in \mathcal{C}$ (how many times $\hat{P}_\alpha$ occurs in $\mathcal{C}$), and denote this multiplicity as $m_\mathcal{C}(\hat{P}_\alpha)$. If one has $m_\mathcal{C}(\hat{P}_\alpha) = |\mathcal{H}_A^{(\alpha)}|/2$, then $\hat{P}_\alpha$ is a normalizer. Hence, with $m_\mathcal{C}(\hat{P}_\alpha)$ and $\mathcal{H}_A^{(\alpha)}$ for all unique $\hat{P}_\alpha \in \mathcal{C}$ computed, one can determine which $\hat{P}_\alpha \in \mathcal{C}$ are normalizers. Further, it can be shown that the multiplicity for $\hat{P}_\alpha$ is upper bounded by $m_\mathcal{C}(\hat{P}_\alpha) \leq |\mathcal{H}_A^{(\alpha)}|/2$, with equality met for normalizers. In any case, the number of new terms introduced to the Hamiltonian upon transformation generated by $\hat{P}_\alpha \in \mathcal{C}$ can be equated to

$$\gamma_\alpha = |\mathcal{H}_A^{(\alpha)}| - 2m_\mathcal{C}(\hat{P}_\alpha), \tag{4.11}$$

where γ_α is the growth incurred by $\hat{P}_\alpha$. In essence, this is the general structure used in finding Pauli products with minimal growth. However, computing $|\mathcal{H}_A^{(\alpha)}|$ for all $\hat{P}_\alpha \in \mathcal{C}$ scales as $O(M^3)$, which can be prohibitively expensive for large problem instances. Further, we generally wish to find $\hat{P}_\alpha$ not only with low γ_α values, but also with high energy gradients g_α. We present below approximate deterministic and probabilistic algorithms for obtaining low growth Pauli products with a fixed energy gradient.

Approximate deterministic algorithm

We aim to find a Pauli product with the lowest γ_α that is in a given non-zero gradient partition. As reviewed in Section 2.2, all Pauli products within a partition are characterized by an "x-string" $\vec{\mu}^{(k)}$ with identical gradient magnitude $g^{(k)}$. We refer to the lowest γ_α found for this set as $\gamma^{(k)}$. The Pauli product corresponding to $\gamma^{(k)}$ is hence the Pauli product with gradient magnitude $g^{(k)}$ possessing the lowest growth. In essence, we formulate an efficient procedure for obtaining $\gamma^{(k)}$ and the corresponding Pauli product given $\vec{\mu}^{(k)}$. Since a Pauli product is uniquely defined by $\vec{\mu}$ and $\vec{\nu}$, this procedure can be seen as obtaining $\vec{\nu}$ which minimizes γ for fixed $\vec{\mu}^{(k)}$. Herein we explain the algorithm in detail, and summarize the steps in Algorithm 1.

Algorithm 1 Deterministic search for low γ_α Pauli products with fixed gradient $g^{(k)}$. Let $\Omega \equiv \{\vec{\mu}^{(i)}\}_i$ be the set of unique x-strings found in terms of $\mathcal{H}$. Let $\mathcal{Z}_i \equiv \{\hat{Z}_l^{(i)}\}_l$ (see Eq. (2.12)). As in the description in the main text, non-unital coefficients on Pauli product commutator expressions are ignored.

Input $\mathcal{H}$, $\vec{\mu}^{(k)}$, r

Output growth minimizing element $\hat{P}_\alpha$ with gradient $g^{(k)}$ and growth $\gamma^{(k)}$

obtain set Ω
initiate $\mathcal{C} = \{\}$
for $\vec{\mu}^{(i)}, \vec{\mu}^{(j)} \in \Omega$, $i < j$, **do**
    if $\vec{\mu}^{(i)} + \vec{\mu}^{(j)} \mod 2 = \vec{\mu}^{(k)}$ **then**
        for $\hat{Z}_l^{(i)} \in \mathcal{Z}_i$, $\hat{Z}_{l'}^{(j)} \in \mathcal{Z}_j$ **do**
            compute $\hat{C} = [\hat{Z}_l^{(i)}\hat{X}_i, \hat{Z}_{l'}^{(j)}\hat{X}_j]$
            if $\hat{C} \neq 0$ **then**
                update $\mathcal{C} \to \mathcal{C} + \{\hat{C}\}$
            end if
        end for
    end if
end for
count occurrences $m_{\mathcal{C}}(\hat{P}_\alpha)$ for each unique $\hat{P}_\alpha \in \mathcal{C}$
for r highest occurring $\hat{P}_\alpha \in \mathcal{C}$ **do**
    compute $|\mathcal{H}_A^{(\alpha)}|$
    compute and save $\gamma_\alpha = |\mathcal{H}_A^{(\alpha)}| - 2m_{\mathcal{C}}(\hat{P}_\alpha)$
end for
select $\hat{P}_{\alpha'}$ with smallest $\gamma_{\alpha'}$. Then $\gamma^{(k)} = \gamma_{\alpha'}$

Firstly, we iterate over pairs $(\vec{\mu}^{(i)}, \vec{\mu}^{(j)})$ formed from the set of unique x-strings in the Hamiltonian, $\{\vec{\mu}^{(i)}\}_i$, and save pairs which satisfy

$$\left(\vec{\mu}^{(i)} + \vec{\mu}^{(j)}\right) \mod 2 = \vec{\mu}^{(k)}. \tag{4.12}$$

In other words, the first step of the algorithm consists of finding factorizations of the x-string of interest in terms of pairs of x-strings found in the Hamiltonian. In practice one observes that $|\{\vec{\mu}^{(i)}\}_i| \ll M$, and hence this procedure, scaling $O(|\{\vec{\mu}^{(i)}\}_i|^2)$, is not considered relatively burdensome. Once we have the pairs $\{(\vec{\mu}^{(i)}, \vec{\mu}^{(j)})\}$ which satisfy Eq. (4.12), for each pair $(\vec{\mu}^{(i)}, \vec{\mu}^{(j)})$, we iterate over the terms in $\sum_l c_l^{(i)}\hat{Z}_l^{(i)}$ and $\sum_{l'} c_{l'}^{(j)}\hat{Z}_{l'}^{(j)}$, i.e. the generalized Ising Hamiltonians multiplying $\hat{X}_i$ and $\hat{X}_j$ respectively in Eq. (2.12). While iterating, we are computing a subset of $\mathcal{C}$ (Eq. (4.10)), the commutator between elements

of $\mathcal{H}$ with x-strings given by $\vec{\mu}^{(i)}$ and $\vec{\mu}^{(j)}$:

$$\mathcal{C}_{ij} = \left\{ [Z_l^{(i)} \hat{X}_i, Z_{l'}^{(j)} \hat{X}_j] \right\}_{l,l'} / \{0\}. \tag{4.13}$$

We compute $\mathcal{C}_{ij}$ for all valid x-string factorizations $\{(\vec{\mu}^{(i)}, \vec{\mu}^{(j)})\}$, and in doing so, obtain sum of multiplicities for any $\hat{P}_\alpha \in \bigcup_{(ij)} \mathcal{C}_{ij}$ as $\sum_{(ij)} m_{\mathcal{C}_{ij}}(\hat{P}_\alpha)$. In principle, by first finding valid factorizations of the x-string of interest, $\vec{\mu}^{(k)}$, we avoid querying the commutator between Hamiltonian elements which will not possess the correct x-string upon commutation, and hence will not generally result in the desired gradient $g^{(k)}$. To obtain the growths γ_k for $\hat{P}_\alpha \in \bigcup_{(ij)} \mathcal{C}_{ij}$, we must compute $|\mathcal{H}_A^{(\alpha)}|$ for all considered $\hat{P}_\alpha$. This step can still be prohibitively expensive for large M, even if $|\bigcup_{(ij)} \mathcal{C}_{ij}| \ll |\mathcal{C}|$. Instead of exhaustively computing $|\mathcal{H}_A^{(\alpha)}|$ for all unique $\hat{P}_\alpha \in \bigcup_{(ij)} \mathcal{C}_{ij}$, we instead only compute $|\mathcal{H}_A^{(\alpha)}|$ for the Pauli products possessing the top r multiplicities. The growth counts are given by $\gamma_\alpha = |\mathcal{H}_A^{(\alpha)}| - 2\sum_{(ij)} m_{\mathcal{C}_{ij}}(\hat{P}_\alpha)$. For instance, setting $r = \lceil \log_2 M \rceil$ results in $O(M \log(M))$ scaling for computing $|\mathcal{H}_A^{(\alpha)}|$ for the top r Pauli products by multiplicity. There is generally a strong negative correlation between multiplicities and growths, and hence one is often able to set $r \ll |\bigcup_{(ij)} \mathcal{C}_{ij}|$ while still successfully obtaining the lowest growth count element, with growth $\gamma^{(k)}$. Empirical evidence of the correlation between multiplicities and the growth incurred by Pauli products can be seen in Figure 4.1, where the multiplicities versus growth for all 2^{N-1} Pauli products from the highest $g^{(k)}$ pool partition have been plotted for several effective Hamiltonians for the CAS(6e, 6o) N_2 model along the GM(1)-iQCC procedure. Notably, for all considered instances, there are no cases where Pauli product $\hat{P}_\alpha$ not found in $\mathcal{C}$ (i.e., has zero multiplicity) exhibits lower Hamiltonian growth than all $\hat{P}_\alpha \in \mathcal{C}$.

4.3 Results

Within this section we benchmark utilization of iQCC effective Hamiltonians obtained with operator selection employing the hybrid importance measure Eq. (4.2) for polyatomic systems exhibiting strong correlation effects. We compare to the canonical sampling choice used in the conventional iQCC methodology of Chapter 2. To recall, 2^{N-1} Pauli products possessing an odd number of $\hat{y}$ operations with identical $\hat{X}_\alpha$ possess identical gradient g_α. The canonical choice of $\hat{Z}_\alpha$ is defined by placing a single $\hat{y}$ operation on the index of the first non-trivial qubit acted on by $\hat{X}_\alpha$, the remaining non-trivial qubit indices are populated with $\hat{x}$ operators. For example, $\hat{y}_1\hat{x}_2\hat{x}_3\hat{x}_4$ is the canonical iQCC operator choice from a partition characterized by $\hat{X}_\alpha = \hat{x}_1\hat{x}_2\hat{x}_3\hat{x}_4$. Below, we outline the procedure for *growth mitigated* iQCC, denoted GM(a)-iQCC, where a denotes the gradient bias utilized in scoring function $s(a)$ (see Eq. (4.2)). We use $s(a)$ with several different gradient bias values for assessment. All iQCC calculations herein use $N_{\text{gen}} = 1$, hence the highest gradient partition is always selected from in the canonical scheme. Explicitly, the employed procedure for obtaining the

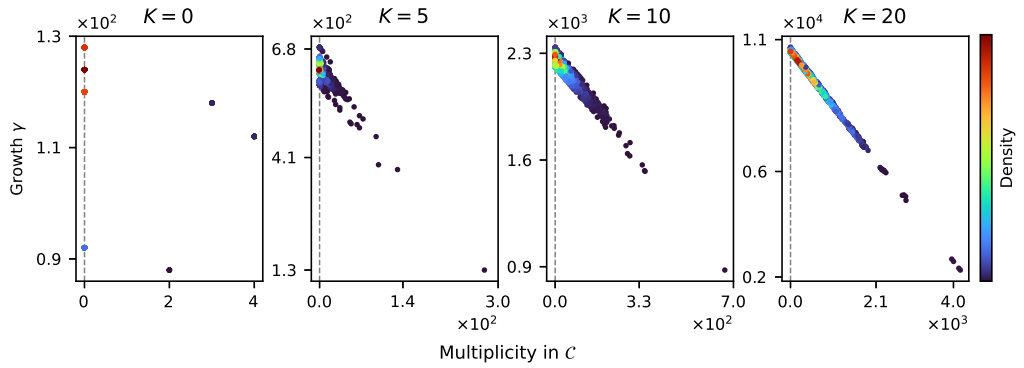

Figure 4.1: Comparison of multiplicity in $\mathcal{C}$ and growth incurred by Pauli products in the top gradient partition for various effective Hamiltonians along the GM(1)-iQCC procedure applied to the CAS(6e, 6o) N_2 system at bond length $R = 1.5$ Å in the cc-pVDZ basis. K denotes the usage of the K^{th} iteration effective Hamiltonian, $\hat{H}^{(K)}$. For each plot, the multiplicities in $\mathcal{C}$ and growths for all $2^{N-1} = 2048$ Pauli products in the top gradient partition of the Pauli pool for $\hat{H}^{(K)}$ are shown. Points in regions of greatest density are coloured red, e.g., $\hat{H}^{(0)}$ only has seven unique data points for all 2048 Pauli products, with a dominating majority having zero multiplicity in $\mathcal{C}$.

K^{th} step effective Hamiltonian in the GM(a)-iQCC scheme is:

1. Obtain DIS description for $\hat{H}^{(K-1)}$. Let $\{g^{(k)}\}_{k=1}^{P}$ be the set of the P highest energetic gradient magnitudes in the DIS, Eq. (2.22).

2. For each of the P highest gradient partitions, obtain the lowest growth incurring element using Algorithm 1 or the techniques described in Appendix D. In doing so, we obtain the minimal growth value of the top P partitions $\{\gamma^{(k)}\}_{k=1}^{P}$.

3. Compute the scores $\{s^{(k)}\}_{k=1}^{P}$ of the top P partitions using chosen bias a, and computed $\{g^{(k)}\}_{k=1}^{P}$ and $\{\gamma^{(k)}\}_{k=1}^{P}$ (see Eq. (4.2)).

4. Select $\hat{P}_K$ corresponding to the growth minimizing element of the top scoring partition according to $\{s^{(k)}\}_{k=1}^{P}$.

5. Perform optimization

$$\min_{\tau_K} E_K(\tau_K) = \langle 0| \, e^{i\tau_K \hat{P}_K/2} \hat{H}^{(K-1)} e^{-i\tau_K \hat{P}_K/2} \, |0\rangle \, . \tag{4.14}$$

   Converged $\min_{\tau_K} E_K(\tau_K)$ is taken as the K^{th} step of the energy estimation.

6. Using $\tau_K^* = \arg\min E_K(\tau_K)$, obtain K^{th} step effective Hamiltonian $\hat{H}^{(K)}$ after standard transformation Eq.(2.24).

It should be emphasized that Step 4 features the selection of the *growth minimizing element* of the highest scoring partition. For instance, the $s(1)$ parameterization results in

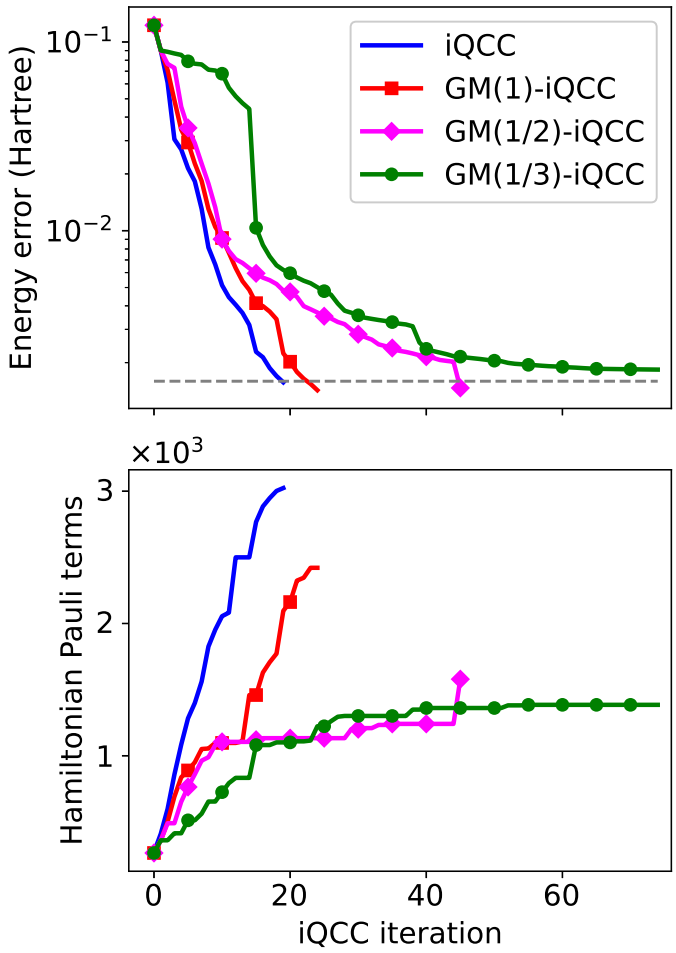

Figure 4.2: Comparison of iQCC procedure convergences applied to the H_4 chain in the STO-3G basis set with an equidistant H−H separation of 1.5 Å. The energy error is the difference between the iQCC energy estimate at a given iteration and the FCI energy. The horizontal grey dashed line indicates an error of 1 kcal/mol (1.6 mHartrees).

the growth minimizing element of the highest energy gradient partition being selected for at each step in the GM(1)-iQCC procedure. This differentiates GM(1)-iQCC from iQCC utilizing the canonical sampling scheme, while also always selects from the highest gradient partition, however with no effort made to minimize growth by the specific Pauli product chosen within the partition.

For all systems studied herein, the JW encoding was utilized in obtaining fermion-to-qubit mapped Hamiltonians with no qubit tapering procedures employed. Hence a fermionic Hamiltonian defined over N_{mo} active molecular orbitals (MOs) results in a qubit-space Hamiltonian of $N = 2N_{\text{mo}}$ qubits. All fermionic Hamiltonians were resolved in the restricted Hartree-Fock (RHF) MO basis. All iQCC calculations utilize a fixed reference state taken to be the qubit-space RHF state, which corresponds to a computational basis state under the mapping of RHF spin orbitals to qubits.

In Figure 4.2, convergence of iQCC energies utilizing various parameterizations of the partition score function are compared with the canonical sampling scheme. The electronic Hamiltonian for the linear H_4 chain was generated in the minimal STO-3G basis set with an equidistant H−H separation of $R = 1.5$ Å. No freezing of orbitals was performed, resulting in a Hamiltonian defined over 4 MOs and hence a $N = 8$ qubit mapped Hamiltonian. Energy errors are taken with respect to the full configuration interaction (FCI) solution, obtained via exact diagonalization. In the GM(a)-iQCC procedures, we use $r = \lceil \log_2 M \rceil$ in the growth minimizing search (see Algorithm 1).

It is observed that operator selection via the $s(1)$ parameterization, utilized in GM(1)-iQCC, gives a similar convergence trajectory to canonical iQCC selection, only requiring a few more iterations to achieve a desired accuracy of 1 kcal/mol (1.6 mHartrees). This is expected, since in both schemes, the highest energy gradient partition is selected from for the current step Hamiltonian. However, while the schemes sample from the same partition for the initial Hamiltonian, this generally changes for later effective Hamiltonians which have different gradient values for the pool partitions. Hence we still expect slight deviations in their energy convergence trajectories. While selection using $s(1)$ demonstrates energetic convergence on par with that obtained using canonical selection, only a decrease of $\sim 20\%$ in the number of Pauli products in the effective Hamiltonians required for achieving an accuracy of 1 kcal/mol is observed. However, comparing the procedures with fixed iteration number, there are several instances where sampling via $s(1)$ provides substantially reduced Pauli product counts for effective Hamiltonians, e.g., $K = 10$, where the GM(1)-iQCC effective Hamiltonian provides a $\sim 50\%$ reduction in the number of Pauli products in the effective Hamiltonian compared to the canonical selection scheme.

More significant reductions in effective Hamiltonian Pauli product counts are obtained using partition score function with non-zero weight placed on growth value. For instance, $s(1/2)$ provides an effective Hamiltonian with $\sim 48\%$ less terms than canonical sampling to achieve an error of less than 1 kcal/mol. However, since in general, the GM(1/2)-iQCC procedure no longer samples from the highest gradient partition at each step, we observe that the rate of energetic convergence is decreased substantially. While offering the most reduced effective Hamiltonians at earlier iterations, the $s(1/3)$ selection fails to systematically converge the iQCC energy beyond an error of 1.84 mHartrees. Further, the GM(1/3)-iQCC energies are significantly higher than the energies obtained via the other selection schemes for the earlier iterations. We attribute the poor convergence of the $s(1/3)$ selection scheme to be a symptom of over-biasing the growth consideration. In this regime, generators are being selected more so by their expected onset of new Pauli products in the next-step effective Hamiltonian, rather than their energy gradients. Since energetic convergence will largely be dictated by the gradients of the selected generators, it is crucial to not under-weight the importance of energetic gradients $g^{(k)}$ in Eq. (4.2).

In Figure 4.3, an assessment of GM-iQCC applied to N_2 at a bond length of $R = 1.5$ Å in

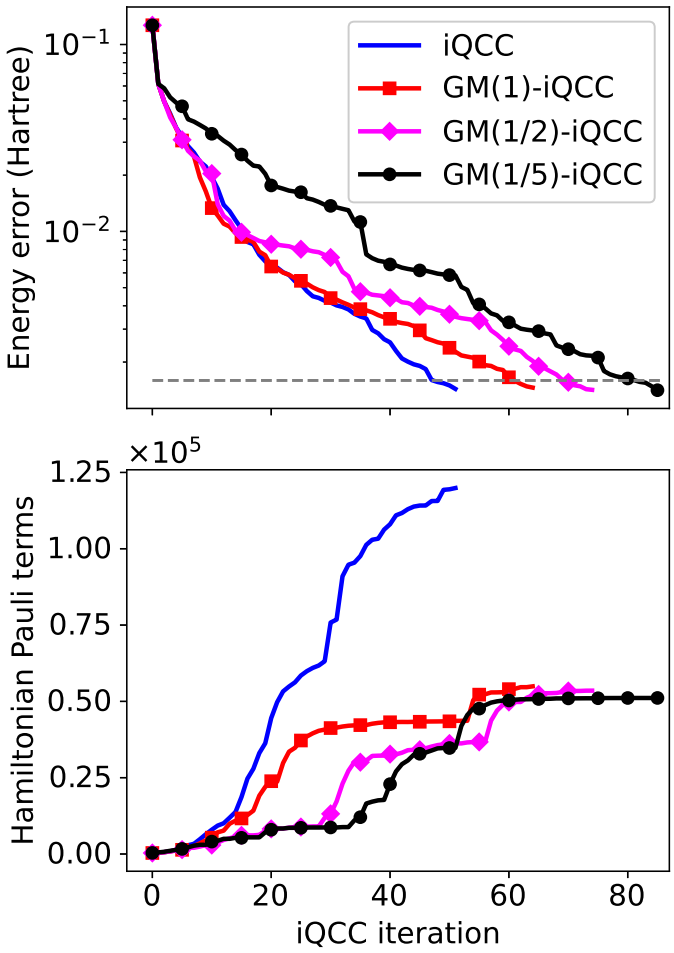

Figure 4.3: Comparison of iQCC procedure convergences applied to a CAS(6e, 6o) N_2 model in the cc-pVDZ basis set with a bond distance of 1.5 Å. The energy error is the difference between the iQCC energy estimate at a given iteration and the CASCI energy. The horizontal grey dashed line indicates an error of 1 kcal/mol (1.6 mHartrees).

an active space derived from the cc-pVDZ basis set is given. The active space is constituted by six electrons in six molecular orbitals (6e,6o), the minimal complete active space to describe the triple bond formation, resulting in a $N = 12$ qubit-mapped Hamiltonian. The three highest energy occupied MOs and three lowest energy unoccupied MOs in the converged RHF solution form the active space. Energy errors are reported relative to the complete active space configuration interaction (CASCI) solution, obtained via exact diagonalization of the active space Hamiltonian. The GM(a)-iQCC procedures utilize $r = \lceil \log_2 M \rceil$ in the growth minimizing search.

The canonical, $s(1)$, and $s(1/2)$ selection schemes give similar energetic convergence for the first ~ 15 iterations. As expected, iQCC and GM(1)-iQCC share similar convergence trajectories, requiring 52 and 65 iterations to achieve an error within 1 kcal/mol, respectively. Significantly, the effective $K = 65$ effective Hamiltonian for GM(1)-iQCC consists of

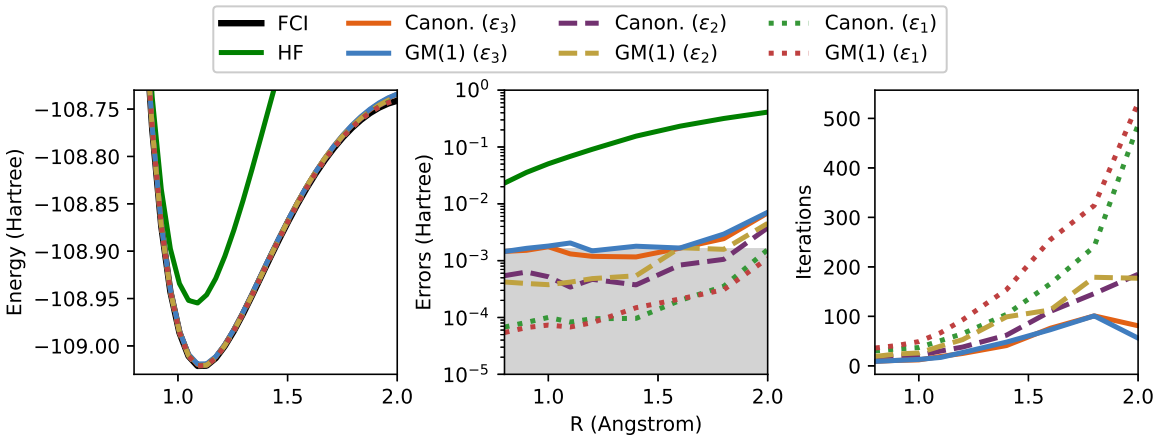

Figure 4.4: Potential energy surfaces obtained by the standard iQCC ("Canon.") and GM(1)-iQCC procedures applied to the CAS(6e, 6o) N_2 model in the cc-pVDZ basis set. The energy error is the difference between the iQCC energy estimate at a given iteration and the CASCI energy. The grey shaded area of the middle plot indicates errors within 1 kcal/mol (1.6 mHartrees) of the exact CASCI energy.

5.5×10^4 terms, a $\sim 46\%$ reduction compared to the $K = 52$ effective iQCC Hamiltonian of 1.2×10^5 terms obtained via canonical selection.

The $s(1/2)$ selection procedure provides drastically reduced term counts for effective Hamiltonians up to the 30^{th} iteration, at which point a rapid onset of terms in the subsequent iterations brings the Pauli product counts to be comparable with the $s(1)$ procedure at similar iterations. By this iteration, all schemes' effective Hamiltonians provide substantially improved reference energies, however the GM(1)-iQCC and GM(1/2)-iQCC procedures offering substantially reduced effective Hamiltonian Pauli product counts. For instance, at $K = 20$, the canonical selection provides an effective Hamiltonian of 3.6×10^4 Pauli products and an energy error of 6.9 mHartrees, while the $s(1/2)$ selection gives an effective Hamiltonian of 7.6×10^3, demonstrating a $\sim 79\%$ reduction, with a comparable energy error of 8.7 mHartrees.

While adaptive ansatz schemes inherently yield non-differentiable energy surfaces, it is desirable that the growth-mitigated iQCC relative errors over the considered bond lengths are not systematically worse than those of the standard iQCC procedure. To this end, potential energy surfaces PESs for the N_2 model obtained via iQCC and GM(1)-iQCC procedures are shown in Figure 4.4. To assess the quality of PESs obtained via GM(1) and standard iQCC procedures at various levels of convergence, we utilize a convergence criterion based on the gradient norm, i.e., the procedure exits when $\sum_k |g^{(k)}| \leq \epsilon_i$ for predefined threshold ϵ_i. Figure 4.4 includes comparison of the GM(1) and standard iQCC procedures at three gradient convergence thresholds: $\epsilon_1 = 0.1$, $\epsilon_2 = 0.2$, and $\epsilon_3 = 0.3$ Hartree.

For fixed convergence threshold, the standard and GM(1) iQCC procedures demonstrate

similar errors and number of required iterations at the majority of bond lengths. Utilizing thresholds ϵ_1, ϵ_2, and ϵ_3, the standard iQCC procedures respectively achieve non-parallelity errors (NPEs) of 1.5×10^{-3}, 3.4×10^{-3}, and 5.7×10^{-3}, while the GM(1) procedures obtain NPEs of 1.0×10^{-3}, 4.1×10^{-3}, and 5.6×10^{-3}. The NPE is defined as the difference between maximum error and minimum error relative to the exact CASCI energy obtained over all considered bond lengths. With considerably similar performance in terms of energetic errors and number of iterations required for convergence, it is to be noted that the advantage of the GM(1) procedure lies in the reduced number of Hamiltonian terms along the iQCC procedure.

In Table 4.1, benchmarking of simulated VQE application of iQCC and GM(a)-iQCC pre-processed Hamiltonians is demonstrated, with comparisons to other classical electronic structure and VQE computations employing the original Hamiltonian. All classical electronic structure calculations were carried out using the `Psi4` quantum chemistry package.[69] The strongly correlated nature of the N_2 model with a bond length of 1.5 Å leads to a challenging case for the post-HF classical electronic structure methods, evident by the energetic errors of the coupled-cluster with singles and doubles (CCSD) and Møller-Plesset second order perturbation theory (MP2) energies being several mHartrees. Similarly, the dissociation of the linear H_4 chain to separated hydrogen atoms is a prototypical model exhibiting strong correlation effects. For the single point energy evaluation used for assessment, we set the equidistant H−H separation as 1.5 Å and utilize the STO-3G minimal basis. Furthermore, benchmarking against the VQE employing the UCCSD ansatz as implemented in the `Tequila` package is performed. The UCCSD ansatz likewise fails to achieve accuracy within 1 kcal/mol for these problem instances. All methods involving VQE computations utilizing iQCC or GM(a)-iQCC preprocessed qubit space Hamiltonian feature transformation including only a single generator per iteration. For example, method "GM(1)-iQCC(10)" in Table 4.1 denotes a preprocessed Hamiltonian obtained after twenty iterations of the iQCC procedure employing the $s(1)$ scoring function for the single generator selection performed at each iteration. Subsequently, the QCC method utilizing ten generators is applied to the final effective Hamiltonian. The energy error reported is the difference between the final optimized QCC energy and the CASCI energy. In general, the iQCC preprocessed Hamiltonians allow for much shallower quantum circuits, at the expense of more Pauli products in the Hamiltonian entering the QCC procedure. The GM(a)-iQCC procedure likewise produces quantum resource reductions competitive with iQCC, while offering substantially reduced effective Hamiltonian Pauli product counts.

As a larger scale example, we compare performance of the first few iterations of iQCC and GM(1)-iQCC for H_2O in the 6−31G(d) basis set. The lowest-energy MO, corresponding to the $1s$ atomic orbital of the oxygen atom, was frozen, with the remaining 18 MOs constituting the active space, yielding a 36 qubit Hamiltonian. A symmetric O−H bond distance of 1.5 Å was used, with H−O−H angle set to 107.60°. The initial qubit space

Table 4.1: Comparison of methods applied to models in the strongly correlated regime, namely CAS(6e, 6o) N_2 in the cc-pVDZ basis set and a linear H_4 chain in the STO-3G basis set. Both the N−N bond distance and the equidistant H−H separation were set to $R = 1.5$ Å, where both systems possess a high degree of orbital degeneracies. The HF, MP2, and CCSD computations were carried out using the `Psi4` package. [69] The UCCSD calculation were performed using the `Tequila` package, where the cluster operator was obtained in closed-shell form before a first-order Trotterization. [70] QCC(N_{ent}) denotes the QCC ansatz constructed from N_{ent} selected generators. (GM(a)-)iQCC(N_{ent}) denotes the QCC(N_{ent}) procedure using (GM(a)-)iQCC preprocessed Hamiltonian. All preprocessed Hamiltonians were obtained after $K = 20$ iterations. Errors are reported in mHartrees. The number of Pauli products in the Hamiltonian utilized in state vector VQE simulation is denoted by M. The reported CNOT counts were obtained via circuit compilation using the `Tequila` package with the `Qiskit` backend. [71]

Method	N_2				H_4			
	Error	M	CNOTs	Circuit parameters	Error	M	CNOTs	Circuit parameters
HF	192	-	-	-	167	-	-	-
CCSD	8.02	-	-	-	−1.47	-	-	-
MP2	−16.8	-	-	-	80.5	-	-	-
UCCSD	2.84	247	4512	30	2.59	185	1280	12
QCC(10)	14.6	247	60	10	3.30	185	60	10
iQCC(10)	4.66	36281	88	10	0.522	3024	80	10
GM(1)-iQCC(10)	4.62	21992	68	10	0.825	2096	84	10
GM($\frac{1}{2}$)-iQCC(10)	7.48	8567	92	10	1.62	1133	72	10
QCC(50)	1.69	247	300	50	0.161	185	252	50
iQCC(50)	0.670	36281	452	50	4.36×10^{-2}	3024	300	50
GM(1)-iQCC(50)	0.814	21992	416	50	5.00×10^{-3}	2096	328	50
GM($\frac{1}{2}$)-iQCC(50)	1.22	8567	436	50	4.65×10^{-3}	1133	316	50

Hamiltonian, obtained via the JW transformation of the fermionic Hamiltonian in the canonical orbital basis, contains 41915 Pauli products. To assess the relative quality of iQCC energies obtained via the canonical and hybrid operator selection scheme, we obtain energetic errors with respect to the CASCI energy, obtained via Davidson diagonalization in the determinantal basis using the `GAMESS` package. [72]

The canonical iQCC and GM(1)-iQCC procedures are compared for H_2O in Table 4.2. In this instance, we use $r = \lceil M/10 \rceil$, as the growth minimizing search using $r = \lceil \log_2 M \rceil$ could not find the growth minimizing element in the highest gradient partition for the initial Hamiltonian. This case can arise when there exists a Pauli product in the searched partition which incurs a low growth as a result from commuting with an anomalously large number of terms in the Hamiltonian. For the first six iterations, the canonical choice and the growth minimizing choice found within the highest gradient partition via the $s(1)$ procedure provide identical optimized trial energies up to 10^{-4} Hartree. The reduced proliferation of new Pauli products in the GM(1)-iQCC procedure becomes pronounced as further iterations are performed. For the eighth iteration, a 43.5 percent reduction in terms using the $s(1)$ selection procedure (2.17×10^5 Pauli products) compared to the canonical selection (3.84×10^5 Pauli products) is observed. Furthermore, the generators selected along the $s(1)$ iQCC procedure also provide a slightly lower optimized energy by the eighth iteration compared to the iQCC procedure employing the canonical selection.

Table 4.2: Comparison of the iQCC and GM(1)-iQCC procedures applied to the 36 qubit H_2O model in the 6−31G(d) basis with symmetric bond length of 1.5 Å. Numerics for the iQCC and GM(1)-iQCC methods for a given iteration are tabulated in uncoloured and grey rows respectively.

Iteration	Energy error (Hartree)	Pauli products
1	0.2319	5.63×10^4
	0.2319	5.63×10^4
2	0.2293	6.79×10^4
	0.2293	6.65×10^4
4	0.2092	9.63×10^4
	0.2092	9.12×10^4
6	0.2048	1.82×10^5
	0.2048	1.41×10^5
8	0.2023	3.84×10^5
	0.2016	2.17×10^5

4.4 Conclusions

In this chapter, we have introduced a Pauli term ranking importance measure which takes into account the expected onset of terms in the iteratively transformed Hamiltonian. The hybrid importance measure, which has contributions from expected growth *and* energetic gradients, was utilized in the selection of iQCC generators for molecular systems in strongly correlated regimes. Due to the exponential size of the DIS, efficient algorithms to finding growth-minimizing elements were devised which avoid exhaustively computing the expected growth for all possible elements. It was found that so long as biasing towards energetic gradients is not made too small in the pool partition score function, we retain robust energetic convergence towards FCI and CASCI energies, while achieving significant reductions in the number of resulting Hamiltonian Pauli products. Energetic convergences when using score function $s(1)$ were on par with the canonical iQCC selection procedure, while the former achieved up to a $\sim 50\%$ reduction in the number of resulting Hamiltonian Pauli products to achieve an energy error within 1 kcal/mol. Placing greater bias towards minimizing growth can substantially reduce the size of early iteration effective Hamiltonians, however such a bias can impede energy lowering at later iterations. We hence consider the GM(1)-iQCC procedure to be the most competitive approach to mitigating Hamiltonian growth while substantially reducing quantum resource requirements. Similar to the canonical iQCC procedure, the Hermitian character and eigenspectrum are formally conserved in the effective Hamiltonians, differentiating this approach from the large majority of efforts in VQE Hamiltonian preprocessing. However, the GM(1)-iQCC procedure offers substantially smaller effective Hamiltonians compared to iQCC, while yielding competitive energies for fixed iteration count.

Chapter 5

Excited state estimation

5.1 Overview

Within this chapter, we introduce extensions to the iQCC formalism which facilitate the simultaneous determination of ground and excited states. Instead of constructively obtaining the unitary which rotates a given reference state, $|\phi_0\rangle$ towards target ground state $|\Psi_0\rangle$, the developments described herein allow for the construction of a single unitary $\hat{U}$, which rotates a set of reference states to a set of eigenstates. To allow for additional flexibility of reference states, we introduce a multireference extension of the iQCC procedure, which allows for the application of reference states beyond a single electronic configuration (computational basis state). Given a set of multiconfigurational reference states $\{|I\rangle\}_{I=0}^{N_s-1}$, the procedure utilizes a state-averaged ensemble approach to adaptively construct the unitary which approximately achieves $\{\hat{U}\,|I\rangle\}_{I=0}^{N_s-1} = \{\Psi_i\}_{i=0}^{N_s-1}$.

5.2 State-averaged iQCC

Herein we describe the multi-reference state-averaged (MRSA-) iQCC procedure. As an input to the MRSA-iQCC procedure, one must define a set of orthonormal reference states $\{|I\rangle\}_{I=0}^{N_\mathrm{s}-1}$ to be used as the starting points in the estimation of N_s electronic eigenstates. Herein, $|I\rangle$ are not restricted to be single electronic configuration. Rather, $\{|I\rangle\}_{I=0}^{N_\mathrm{s}-1}$ can be arbitrary linear combination of electronic configurations in a fixed orbital basis. Given a set $\{|I\rangle\}_{I=0}^{N_\mathrm{s}-1}$, we denote the *model space* $\mathcal{P} = \{|\phi_j\rangle\}_{j=1}^{L}$ as the set of electronic configurations with non-zero component in any $|I\rangle$, i.e.,

$$|I\rangle = \sum_{\phi_j \in \mathcal{P}} c_j^{(I)}\,|\phi_j\rangle\,, \tag{5.1}$$

The principal quantity in the MRSA-iQCC procedure is the density for the state-averaged mixture of $\{|I\rangle\}_{I=0}^{N_\mathrm{s}-1}$, given by

$$\rho = \sum_{I=0}^{N_\mathrm{s}-1} w_I \, |I\rangle \langle I| \,, \tag{5.2}$$

where w_I are fixed weights respecting $0 \leq w_I \leq 1$ and $\sum_I w_I = 1$, leading to the state-averaged energy at iteration K:

$$E_{\mathrm{SA}}^{(K)}(\boldsymbol{\tau}) = \mathrm{Tr}\Big(\hat{H}^{(K)}\hat{U}(\boldsymbol{\tau})\hat{\rho}\hat{U}^\dagger(\boldsymbol{\tau})\Big) = \sum_{I=1}^{N_\mathrm{s}} w_I \, \langle I| \, \hat{U}^\dagger(\boldsymbol{\tau})\hat{H}^{(K)}\hat{U}(\boldsymbol{\tau}) \, |I\rangle \,. \tag{5.3}$$

As in the ground state iQCC algorithm, the unitary $\hat{U}(\boldsymbol{\tau})$ is constructed as the product of N_g singly exponentiated Pauli terms,

$$\hat{U}(\boldsymbol{\tau}) = \prod_{\alpha=1}^{N_g} \exp\Big(-i\tau_\alpha \hat{P}_\alpha\Big). \tag{5.4}$$

In order to screen for $\hat{P}_\alpha$ which will have significant capability to variationally lower the current step trial energy, we again rely on utilizing the magnitude of energy derivatives,

$$g_\alpha = \left| \frac{\partial E_{SA}^{(K)}}{\partial \tau_\alpha}\Big|_{\tau_\alpha=0} \right|, \tag{5.5}$$

where

$$\frac{\partial E_{\mathrm{SA}}^{(K)}}{\partial \tau_\alpha}\bigg|_{\tau_\alpha=0} = \sum_{I=1}^{N_s} w_I \frac{\partial}{\partial \tau_\alpha} \, \langle I| \, e^{i\tau_\alpha \hat{P}_\alpha/2}\hat{H}^{(K)}e^{-i\tau_\alpha \hat{P}_\alpha/2} \, |I\rangle \, \Big|_{\tau_\alpha=0}. \tag{5.6}$$

In comparison to the gradient screening process used in the single-state iQCC procedure (as described in Section 2.2), efficiently obtaining $\hat{P}_\alpha$ of maximal g_α in the MRSA case is more involved. We provide a derivation of the generator screening process in Appendix E. Essentially, one can arrive at an expression for g_α of the form

$$g_\alpha = |\,\mathrm{Im}(\theta_\alpha)| \left| \sum_{j=1}^{L} \lambda_j^{(\alpha)} \Xi_j^{(\alpha)} \right|, \tag{5.7}$$

where $\Xi_j^{(\alpha)}$ is a classically computable quantity which depends only on the $\hat{X}_\alpha$ part of candidate $\hat{P}_\alpha$ under factorization $\hat{P}_\alpha = \theta_\alpha \hat{X}_\alpha \hat{Z}_\alpha$, and $\lambda_j^{(\alpha)} = \langle \phi_j | \, \hat{Z}_\alpha \, |\phi_j\rangle$. Similar to the DIS construction of Section 2.2, we can efficiently obtain the set of all possible $\hat{X}_\alpha$'s which *potentially* lead to non-zero g_α. The additional complexity arises from the relative phases decided by choice of $\hat{Z}_\alpha$ through $\{\lambda_j^{(\alpha)}\}_{j=1}^{L}$. Given fixed $\hat{X}_\alpha$ component of $\hat{P}_\alpha$, we refer to

the process of maximizing g_α via variation in the $\hat{Z}_\alpha$ component as the phase alignment problem, and discuss strategies for solving it in Appendix F.

Once the set of N_g generators have been selected for inclusion in the current step $\hat{U}(\boldsymbol{\tau})$, the state averaged energy E_{SA} [Eq.(5.3)] is optimized with respect to $\boldsymbol{\tau}$. This can be carried out via classical compute, or by performing the N_s distinct quantum expectation value estimations in Eq.(5.3) if knowledge on how to prepare $\{|I\rangle\}_{I=0}^{N_s-1}$ is possessed. Following the optimization of amplitudes $\boldsymbol{\tau}$, the next-step Hamiltonian $\hat{H}^{(K+1)}$ is obtained via conjugation of $\hat{H}^{(K)}$ with the optimized $\hat{U}(\boldsymbol{\tau}^*)$, as is done in the ground-state iQCC method, see Eq.(2.24). However, the MRSA-iQCC includes an additional step before an iteration is completed. In the case of homogeneous weights in Eq.(5.3), the state-averaged energy is invariant to unitary rotations within the subspace generated by the references $\{|I\rangle\}_{I=0}^{N_s-1}$. In order to access accurate state-specific energies, the final step of an MRSA-iQCC iteration involves the diagonalization of the current-step Hamiltonian in the subspace of $\{|I\rangle\}_{I=0}^{N_s-1}$, i.e., solving

$$H_{\text{MRSA}}^{(K+1)} \boldsymbol{C}_I = E_I \boldsymbol{C}_I \tag{5.8}$$

where

$$\left[H_{\text{MRSA}}^{(K+1)}\right]_{IJ} = \langle I| \hat{H}^{(K+1)} |J\rangle, \quad I, J = 0, \ldots, N_s - 1 \tag{5.9}$$

The eigenvalues $\{E_I\}_{I=0}^{N_s-1}$ are taken to be the state-specific trial energies of iteration K, and eigenvectors $\{\boldsymbol{C}_I\}_{I=0}^{N_s-1}$ can be used to update the reference states for the $(K+1)^{\text{th}}$ iteration. The processes of 1) generator selection, 2) state-averaged optimization, 3) Hamiltonian dressing, and 4) subspace diagonalization are iterated until convergence is observed in E_{SA}.

5.3 Results

In Figure 5.1, MRSA-iQCC is applied to multiple state determination for the linear equidistant H_4 chain in the STO-3G basis set. We assess the algorithm in determining the four lowest lying states at a stretched geometry, with equidistant H−H separation of $R = 1.9$ Å, roughly twice the separation distance at equilibrium geometry along this reaction coordinate. The qubit Hamiltonian is obtained through the JW map, resulting in a $N_q = 8$ qubit Hamiltonian. To obtain the initial set of reference states $\{|I\rangle\}_{I=0}^{3}$, the initial Hamiltonian was diagonalized in a 8 dimensional subspace defined by electronic configurations which had significant weights in the four lowest lying solutions obtained via CISD.

We also compare the performance differences for two different compression thresholds, ϵ_c. To recall, the iQCC effective Hamiltonians can be compressed to a precision threshold ϵ_c, where Pauli terms with coefficient magnitude less than ϵ_c are pruned. While the compression of $e_c = 10^{-8}$ a.u. achieves all four trial states within chemical accuracy from their target

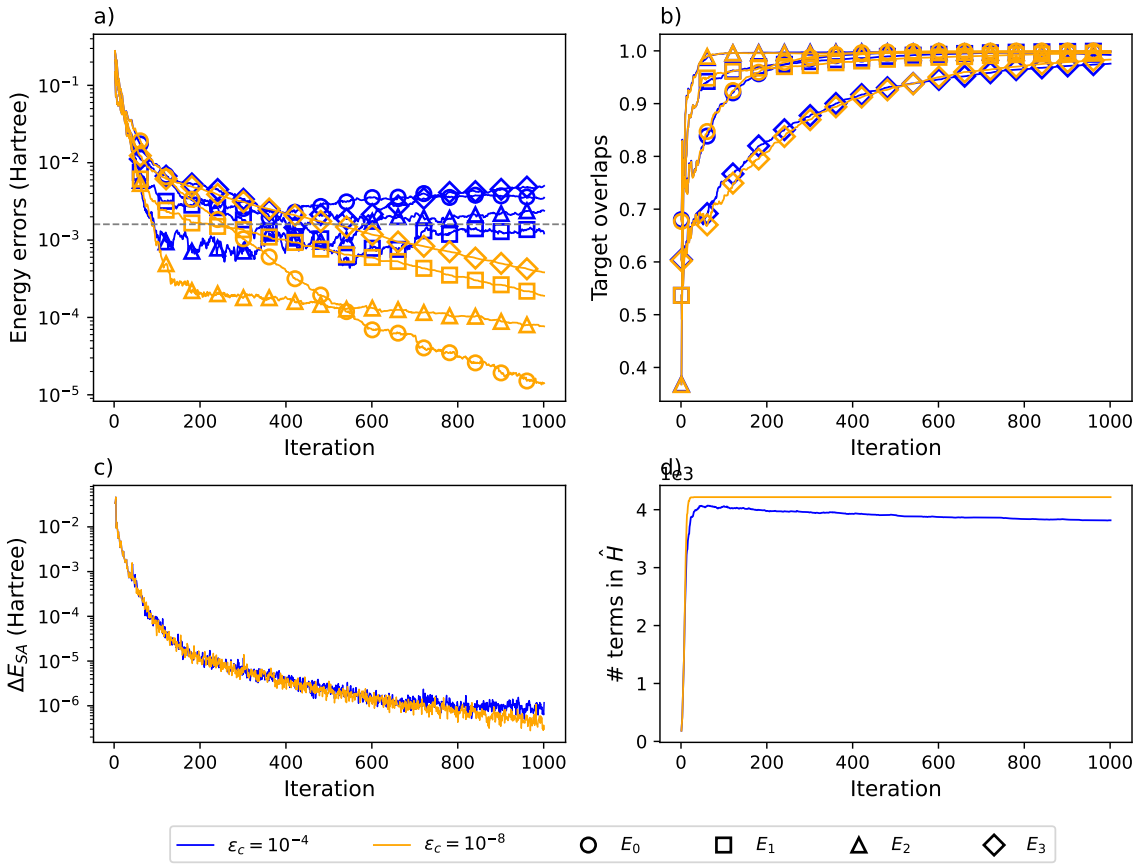

Figure 5.1: MRSA-iQCC applied to the determination of the four lowest energy eigenstates for the H_4 chain with equidistant separation of $2r_e$ in the STO-3G basis set. A total of $L = 8$ computational states were used to define the reference space. The MRSA-iQCC algorithm used $N_g = 1$ generators at each iteration, and utilized the optimal strategy for phase alignment. The ϵ_c denotes the compression threshold used.

states by $\sim$ 500 iterations utilizing a single generator per iteration, the more aggressive compression threshold of $e_c = 10^{-4}$ a.u. fails to achieve chemical accuracy for all four states by the 1000^{th} iteration. Despite the poor energetic convergence under aggressive compression, the target overlaps obtained are on par with those obtained using $e_c = 10^{-8}$.

In Figures 5.2 and 5.3, we compare two of the phase alignment strategies (described in Appendix F) for the problem of simultaneous estimation of the two lowest lying singlet states (S_0, S_1) of a stretched water molecule. Both calculations utilize $N_g = 1$, with standard compression of 10^{-8}. In Figure 5.3, it is seen that both S_0 and S_1 trial energies are readily driven towards their exact solution, approaching machine precision in energy errors using the optimal phase alignment. In contrast, the `GreedySAT` subroutine achieves chemical accuracy with around $\sim$ 50 more iterations, and the convergence to exact energies is notably slower at the late-stage iterations. This can be attributed to the fact that the `GreedySAT` is often missing the true maximal gradient $\hat{P}_\alpha$, and instead is yielding a generator

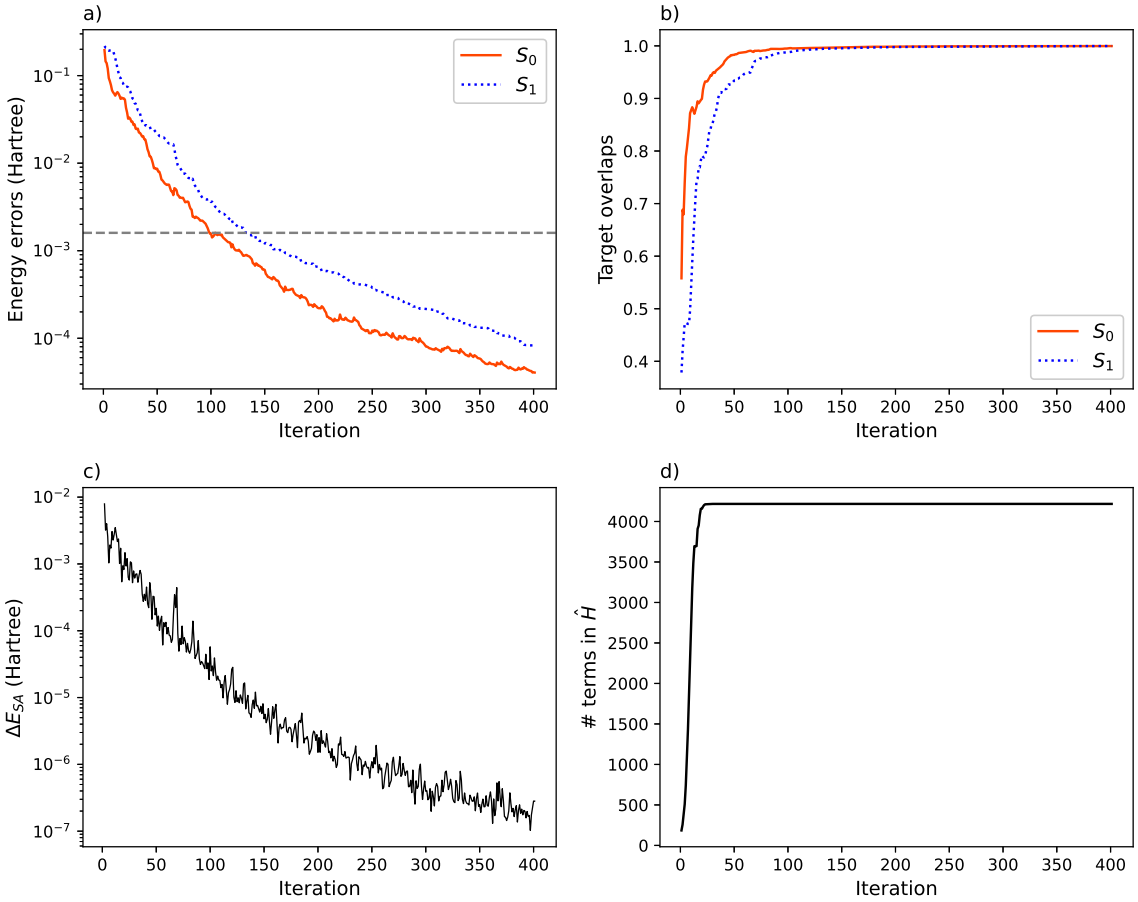

Figure 5.2: Utilization of the `GreedySAT` phase alignment strategy applied to the 4e,4o active space of H_2O in the 6-31G basis set with symmetric O−H bond length of 2.35 Å.

of non-zero, yet suboptimal gradient. The trajectory of the trial S_0 and S_1 overlaps are nearly identical using the two phase alignment procedures.

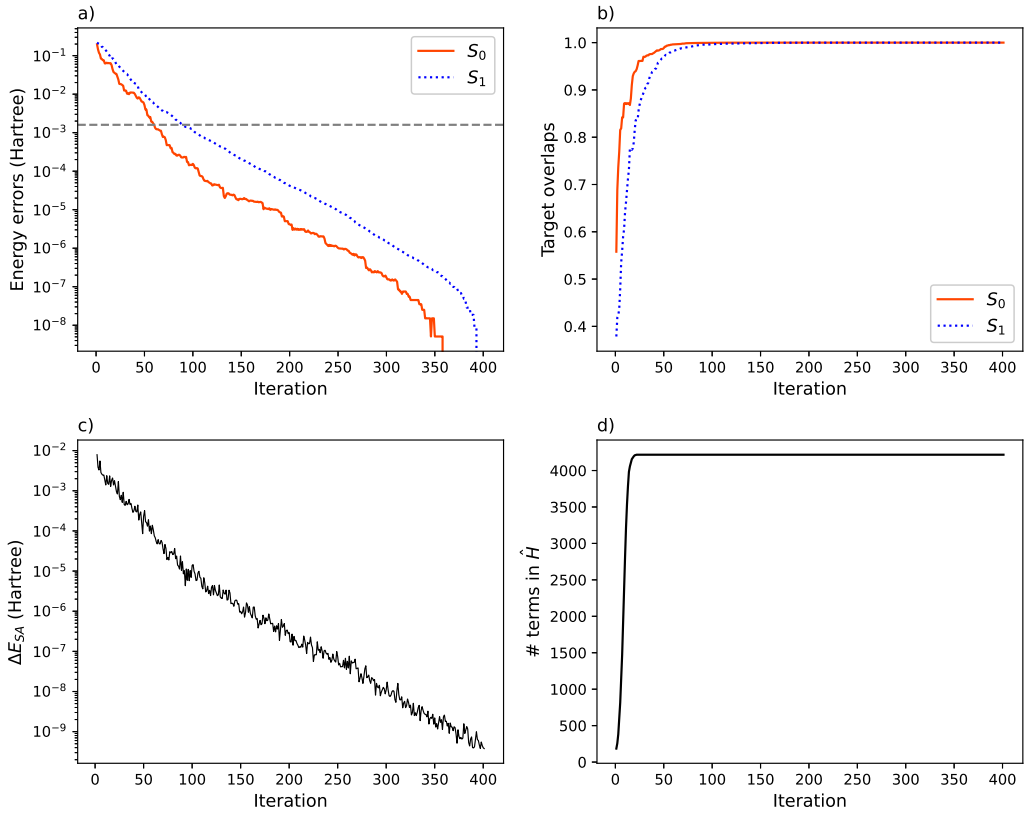

Figure 5.3: Same as Figure 5.2 but utilizing the optimal phase alignment routine.

5.4 Conclusions

This chapter presented presented generalization of the ground state iQCC algorithm, which enables simultaneuous and unbiased determination of ground and excited states. Utilizing a state-average over reference states for generator screening and energy optimization, we devised a screening procedure for finding $\hat{P}_\alpha$ of maximum gradient. However, due to the introduction of relative phase factors, the $\hat{Z}_\alpha$ component of $\hat{P}_\alpha$ is no longer generally a gauge; it can significantly alter the state-averaged gradient. This prompted the formulation for finding $\hat{Z}_\alpha$ which maximizes g_α for fixed $\hat{X}_\alpha$ component. This procedure is generally exponentially hard finding in the general case, hence we must rely on heuristic phase alignment algorithms. It was reported that the MRSA-iQCC algorithm provides target overlaps which are robust even under aggressive compression thresholds. This positions the MRSA-iQCC algorithm as a potentially promising classical state preparation routine for early fault-tolerant QPE.

Chapter 6

Conclusions and outlook

This thesis has described efforts made to address the electronic structure problem using the iQCC method. The method, while originally proposed as a near-term quantum algorithm, has effective applicability as a fully classical "quantum-inspired" electronic structure method under certain algorithmic parameters. The classically implemented procedure also naturally positions itself as a zero quantum overhead state-preparation routine for QPE in an early fault-tolerant quantum setting, due to the ability to produce extremely compact circuit parameterizations of ground states via the obtained iQCC unitary transformation. The algorithm was presented in detail in Chapter 2. It was described how the Hamiltonian itself drives the iterative construction of the ground state parameterization. The iQCC algorithm readily identifies basis generators of the full N_q-qubit Lie algebra which are guaranteed to variationally lower the trial energy. Due to the Hamiltonian self-selecting energetically important generators from a totally unrestricted operator pool, the iQCC algorithm can be considered a bonafide black-box electronic structure method. The main computational bottleneck of the algorithm is the proliferation of terms in the iQCC effective Hamiltonians. While formally exponentially scaling, it has been noted that numerical compressions can significantly reduce the total number of terms and even produce empirical asymptotic scalings much lower than the theoretical exponential estimate.[5]

Nevertheless, since virtually all parts of the algorithm scale at least linearly with the number of terms in the effective Hamiltonian, we presented two algebraic approaches to address the escalating number of terms. We discussed the application of iQCC unitaries involving *involutory linear combinations* (ILCs) of Pauli terms in Chapter 3. Such ILCs can be utilized to generate unitary operations involving N_g Pauli terms while only inducing a multiplicative Hamiltonian growth factor of $O(N_g^2)$. Such a growth factor is in high contrast to the formal growth factor of the standard iQCC procedure, which is $O((3/2)^{N_g})$, and iQCC-ILC was seen to produce significantly smaller effective Hamiltonians for sufficiently large problems. However, repeating the process of Hamiltonian transformation with the iQCC-ILC unitary K times inevitably leads to $O(N_g^{2K})$ scaling. Chapter 4 describes an

alternative approach to mitigating the onset of new terms in the effective Hamiltonians, referred to as the growth-mitigated iQCC (GM-iQCC) method. In this approach, Pauli terms are sampled from the N_q-qubit Lie algebra which are energetically important while also minimizing the number of expected terms introduced in the updated Hamiltonian. While this approach does not seem to offer a formal complexity reduction of the transformed Hamiltonians, generally a reduction of $\sim 50\%$ was achieved in the number of terms when compared to the standard iQCC iteration. It is an essentially unexplored topic how to best use this approach in tandem with the iQCC-ILC approach of Chapter 3. It is a seemingly reasonable approach to perform a single iteration of iQCC-ILC and then switch to searching for growth-minimizing generators using GM-iQCC, considering the iQCC-ILC unitary may be best suited for when amplitudes are expected to be high. In this case, one cannot count on numerical compressions to appreciably reduce the number of terms. Furthermore, typically GM-iQCC can only find generators which incur small Hamiltonian growth after the first few iterations, and hence it generally has limited utility on the initial Hamiltonian.

In Chapter 5, we provided a generalization of the iQCC algorithm which allows for the simultaneuous determination of electronic ground and excited states by means of the construction of a single unitary transformation. The technique takes a state-averaging approach to systematically rotating the set of reference states towards their target states. While convergence of all trial energies to within chemical accuracy of their target eigenstates can take a substantial number of iterations, depending on various problem parameters, it was noted that generally FCI overlaps see an appreciable increase after a modest number of iterations. Hence, this approach is particularly promising as a classical ground and excited state-preparation algorithm for QPE in an early fault-tolerant setting.

In Appendix A, formal aspects of the iQCC algorithm are analyzed on the footing of fermionic algebra. It is established that arbitrary fermionic excitations $\hat{E}^{ab\ldots c}_{ij\ldots k}$ can be represented in the form $\hat{E}^{ab\ldots c}_{ij\ldots k} = \hat{\Gamma}\hat{D}$, where $\hat{\Gamma}$ is a Majorana string, and $\hat{D}$ is a projector. By considering the isomorphism between Majorana strings and Pauli terms through fermion-to-qubit mappings, it's found that the iQCC generators correspond to the qubit image of $\hat{\Gamma}$ components of non-zero energy gradient fermionic excitations. While the generator screening procedure can only identify $\hat{\Gamma}$'s arising from double excitations for the initial Hamiltonian, sufficiently iterating the iQCC Hamiltonian transformation allows for the identification of $\hat{\Gamma}$'s originating from *arbitrary rank* fermionic excitations. This fact theoretically supports the robust convergence of the iQCC method, and highlights differences between itself and other schemes, such as ADAPT-VQE,[45] or (generalized) UCC-based ansätze,[14, 34–39] formulated in terms of fermionic excitations of fixed rank. In light of the analysis in fermionic algebra, it is more-or-less appropriate to classify the iQCC algo-

rithm as a "symmetry broken variational selected CC" method, where the CC evolution operator is constructively built by selecting excitations expected to have largest capability for lowering the energy.

Since it may take some time before a practical quantum advantage for electronic structure is definitively realized,[73–75] pursuing the avenue of quantum-inspired classical algorithms for electronic structure can lead to potentially fruitful new insights. Formulating the electronic structure problem in qubit algebra can result in potentially exotic paths towards the development of new methods, either for application of near-term or future quantum resources, or immediately applicable classical algorithms.

Appendix A

iQCC in fermionic algebra

Due to its formulation in the qubit-space algebra, aspects of the iQCC methodology are less than straightforward to formally compare to other electronic structure techniques. By considering the iQCC ansatz in the fermionic algebra, this Appendix attempts to better connect aspects of the iQCC algorithm to the electronic structure literature.

A.1 Generator representation as Majorana strings

The fermionic creation ($\hat{a}_p^\dagger$) and annihilation ($\hat{a}_p$) operators are related to the fermionic Majorana operators through

$$\hat{\gamma}_{2p} = \hat{a}_p + \hat{a}_p^\dagger \tag{A.1}$$

$$\hat{\gamma}_{2p+1} = -i\left(\hat{a}_p - \hat{a}_p^\dagger\right). \tag{A.2}$$

For a fermionic Fock space of M modes, there are $2M$ Majorana operators, and they satisfy

$$\{\hat{\gamma}_i, \hat{\gamma}_j\} = 2\delta_{ij}. \tag{A.3}$$

Labelling the Fock modes $(0, 2, \ldots, M-1)$, and defining $\mathcal{M}$ as the vector space with basis $\{\hat{\gamma}_i\}_{i=0}^{2M-1}$, one obtains the Clifford algebra $Cl(\mathcal{M})$ taking Eq.(A.3) as the quadratic form, with dimensionality $\dim(Cl(\mathcal{M})) = 2^{2M} = 4^M$. The basis elements of $Cl(\mathcal{M})$ are

$$\left\{\hat{\Gamma}_{\boldsymbol{s}} = \hat{\gamma}_0^{s_0}\hat{\gamma}_1^{s_1}\ldots\hat{\gamma}_{2M-2}^{s_{2M-2}}\hat{\gamma}_{2M-1}^{s_{2M-1}} : \boldsymbol{s} = (s_0, s_1, \ldots s_{2M-2}, s_{2M-1}) \in \{0,1\}^{\otimes 2M}\right\}, \tag{A.4}$$

where we have indexing of binary vector $\boldsymbol{s}$ starting at 0 for consistency with basis of $\mathcal{M}$. We refer to $\Gamma_{\boldsymbol{s}}$ as a Majorana *string*. Arbitrary $\Gamma_{\boldsymbol{s}}$ are unitary, as well as involutory up to multiplication by imaginary unit $\pm i$, i.e., $(\Gamma_{\boldsymbol{s}})^2 \in \{-1, 1\}$.* Up to unit phases, Eq.(A.4) is

*This is because a Majorana string is either Hermitian or anti-Hermitian, depending on its degree (the number of Majorana operators it possesses).

isomorphic to the set of M-qubit Pauli terms including identity, $\{\hat{P}_\alpha\}_{\alpha=1}^{4^M}$,

$$\upsilon_{\boldsymbol{s}}\hat{\Gamma}_{\boldsymbol{s}} \cong \hat{P}_{\boldsymbol{s}}, \quad \upsilon_{\boldsymbol{s}} \in \{1, -1, i, -i\}. \tag{A.5}$$

The fermion-to-qubit maps such as JW establish such isomorphisms: any Majorana string is mapped uniquelly to a single Pauli term. We denote the Pauli term arising from $\hat{\Gamma}_{\boldsymbol{s}}$ as $\hat{P}_{\boldsymbol{s}}$ via the chosen fermion-to-qubit mapping.

Therefore, by considering the inverse of the fermion-to-qubit map, the wavefunction ansatz of the iQCC method has the general form in terms of Majorana strings given by

$$\hat{U}(\boldsymbol{\tau})\,|\phi_0\rangle = \prod_{\alpha=1}^{N_g} \exp\Big(-i\tau_{\boldsymbol{s}}\hat{P}_{\boldsymbol{s}}/2\Big)\,|\phi_0\rangle \tag{A.6}$$

$$\cong \prod_{\alpha=1}^{N_g} \exp\left(\frac{-i\tau_{\boldsymbol{s}}\upsilon_{\boldsymbol{s}}\hat{\Gamma}_{\boldsymbol{s}^{(\alpha)}}}{2}\right)|\phi_{0,f}\rangle\,, \tag{A.7}$$

where $|\phi_{0,f}\rangle$ denotes a Slater determinant formally living in M-mode Fock space, as opposed to computational basis state $|\phi_0\rangle$ living in the M-qubit Hilbert space.

To better illustrate the nature of Eq.(A.7), it is useful to put fermionic excitations on the same footing as Majorana strings. As the simplest example, a single excitation may be written as

$$\hat{a}_a^\dagger\hat{a}_i = \frac{1}{4}\left(\hat{\gamma}_{2a}\hat{\gamma}_{2i} + i\hat{\gamma}_{2a}\hat{\gamma}_{2i+1} - i\hat{\gamma}_{2a+1}\hat{\gamma}_{2i} + \hat{\gamma}_{2a+1}\hat{\gamma}_{2i+1}\right) \tag{A.8}$$

which can be recast as

$$\hat{a}_a^\dagger\hat{a}_i = \frac{1}{4}\hat{\gamma}_{2a}\hat{\gamma}_{2i}\left(\hat{1} + i\hat{\gamma}_{2i}\hat{\gamma}_{2i+1} - i\hat{\gamma}_{2a}\hat{\gamma}_{2a+1} + \hat{\gamma}_{2a}\hat{\gamma}_{2a+1}\hat{\gamma}_{2i}\hat{\gamma}_{2i+1}\right) \tag{A.9}$$

$$= \hat{\gamma}_{2a}\hat{\gamma}_{2i}\cdot D(\{\hat{\gamma}_{2p}\hat{\gamma}_{2p+1}\}), \tag{A.10}$$

where we have introduced $D(\{\hat{\gamma}_{2p}\hat{\gamma}_{2p+1}\})$ as a general notation for polynomials of elements of the form $\hat{\gamma}_{2p}\hat{\gamma}_{2p+1}$ arising from factorization Eq.(A.10). It can be shown that such $D(\{\hat{\gamma}_{2p}\hat{\gamma}_{2p+1}\})$ are idempotent, $D(\{\hat{\gamma}_{2p}\hat{\gamma}_{2p+1}\})^2 = D(\{\hat{\gamma}_{2p}\hat{\gamma}_{2p+1}\})$. Such terms can be connected to occupation number operators $\hat{n}_p = \hat{a}_p^\dagger\hat{a}_p$,

$$\hat{\gamma}_{2p}\hat{\gamma}_{2p+1} = -i\left(2\hat{n}_p - 1\right). \tag{A.11}$$

Relatedly, any $D(\{\hat{\gamma}_{2p}\hat{\gamma}_{2p+1}\})$ is mapped to a polynomial of strictly $\hat{z}_p$ operations by the fermion-to-qubit mappings. In the JW mapping, $-i(2\hat{n}_p - 1) = i\hat{z}_p$. The form of $\hat{a}_a^\dagger\hat{a}_i$ given

in Eq.(A.10) can be generalized to arbitrary rank excitations $\hat{E}^{ab\ldots c}_{ij\ldots k} \equiv \hat{a}^\dagger_a \hat{a}^\dagger_b \ldots \hat{a}^\dagger_c \hat{a}_i \hat{a}_j \ldots \hat{a}_k$:

$$\hat{E}^{ab\ldots c}_{ij\ldots k} = \hat{\gamma}_{2a}\hat{\gamma}_{2b}\ldots\hat{\gamma}_{2c}\hat{\gamma}_{2i}\hat{\gamma}_{2j}\ldots\hat{\gamma}_{2k} \cdot D(\{\hat{\gamma}_{2p}\hat{\gamma}_{2p+1}\}) \tag{A.12}$$

$$= \Gamma_{\boldsymbol{e}} D(\{\hat{\gamma}_{2p}\hat{\gamma}_{2p+1}\}), \tag{A.13}$$

where $\boldsymbol{e}$ is the binary vector with $e_{2p} = 1$ for $p \in \{a, b, \ldots, c, i, j, \ldots k\}$, and all other components zero. The projector for general rank r has the form $D(\{\hat{\gamma}_{2p}\hat{\gamma}_{2p+1}\}) = (-1)^r \hat{\Gamma}_{\boldsymbol{e}} \hat{E}^{a,b,\ldots c}_{i,j,\ldots k}$, and we provide a proof of its idempotency in Section A.4.

The rest of the discussion will consider arbitrary excitation $\hat{E}^{ab,\ldots c}_{ij\ldots k}$, unless special cases are raised. Hence to ease notation we refer to arbitrary excitation of $|\phi_{0,f}\rangle$ as $\hat{E}$, with Majorana form $\hat{\Gamma}_{\boldsymbol{e}} D(\{\hat{\gamma}_{2p}\hat{\gamma}_{2p+1}\})$. Multiplication of $\Gamma_{\boldsymbol{e}}$ by terms in $D(\{\hat{\gamma}_{2p}\hat{\gamma}_{2p+1}\})$ simply amounts to substitutions of even indices $2p$ for odd indices $2p+1$ in $\Gamma_{\boldsymbol{s}}$ in Eq.(A.12). Hence, while $D(\{\hat{\gamma}_{2p}\hat{\gamma}_{2p+1}\})$ is a diagonal inhomogeneous polynomial, multiplication by $\Gamma_{\boldsymbol{s}}$ in Eq.(A.12) always yields an expansion of Majorana strings which is a homogeneous polynomial: a rank r fermionic excitation yields an expansion of 2^r Majorana strings of degree $2r$, meaning Majorana strings $\Gamma_{\boldsymbol{s}}$ with $|\boldsymbol{s}| = 2r$. Essentially, we can factorize an arbitrary excitation $\hat{E}$ as a single non-diagonal Majorana string $\Gamma_{\boldsymbol{e}}$ multiplying 2^r diagonal Majorana strings in $D(\{\hat{\gamma}_{2p}\hat{\gamma}_{2p+1}\})$.

We are now in a position to connect the Majorana strings identified to have non-zero energy gradient in the iQCC generator screening procedure to fermionic excitations. We can consider the energy derivative with respect to state

$$e^{\tau\hat{E}/2}\,|\phi_0, f\rangle\,, \tag{A.14}$$

i.e., a CC-like wavefunction with only one excitation included in the cluster operator. The energy derivative at $\tau = 0$ is

$$\frac{\partial E}{\partial \tau}\Big|_{\tau=0} = \frac{\partial}{\partial \tau}\,\langle\phi_{0,f}|\,e^{\tau\hat{E}^\dagger/2}\hat{H}e^{\tau\hat{E}/2}\,|\phi_{0,f}\rangle\,\Big|_{\tau=0} \tag{A.15}$$

$$= \frac{1}{2}\,\langle\phi_{0,f}|\,\hat{E}^\dagger\hat{H}\,|\phi_{0,f}\rangle + \frac{1}{2}\,\langle\phi_{0,f}|\,\hat{H}\hat{E}\,|\phi_{0,f}\rangle \tag{A.16}$$

$$= \mathrm{Re}\,\langle\phi_{0,f}|\,\hat{H}\hat{E}\,|\phi_{0,f}\rangle \tag{A.17}$$

$$= \langle\phi_{0,f}|\,\hat{H}\hat{E}\,|\phi_{0,f}\rangle\,, \tag{A.18}$$

where the final equality holds under our typical assumptions of real $\hat{H}$ and real reference function $|\phi_{0,f}\rangle$. We then utilize the decomposition of $\hat{E}$ in Eq.(A.13),

$$\frac{\partial E}{\partial \tau}\Big|_{\tau=0} = \langle\phi_{0,f}|\,\hat{H}\hat{\Gamma}_{\boldsymbol{s}} D(\{\hat{\gamma}_{2p}\hat{\gamma}_{2p+1}\})\,|\phi_{0,f}\rangle \tag{A.19}$$

$$= \langle\phi_{0,f}|\,\hat{H}\hat{\Gamma}_{\boldsymbol{s}}\,|\phi_{0,f}\rangle\,, \tag{A.20}$$

where the second equality holds since $D(\{\hat{\gamma}_{2p}\hat{\gamma}_{2p+1}\})$ projects out the kernel of $\hat{E}$, which

by definition, does not contain configuration $|\phi_{0,f}\rangle$ (arbitrary $\hat{E}^{a,b,\dots c}_{i,j,\dots k}$ is defined such that modes $\{a,b,\dots c\}$ and $\{i,j,\dots k\}$ are respectively empty and occupied in $|\phi_{0,f}\rangle$). Utilizing $\hat{P}_{\boldsymbol{e}} \cong \upsilon_{\boldsymbol{e}}\hat{\Gamma}_{\boldsymbol{e}}$ the iQCC energy derivatives [Eq.(2.9)] have the form

$$\frac{\partial E}{\partial \tau_{\boldsymbol{e}}}\Big|_{\tau_{\boldsymbol{e}}=0} = \mathrm{Im}\left(\upsilon_{\boldsymbol{e}} \langle\phi_0| \hat{H}\hat{\Gamma}_{\boldsymbol{e}} |\phi_0\rangle\right), \tag{A.21}$$

and therefore, for purely imaginary $\upsilon_{\boldsymbol{e}}\hat{\Gamma}_{\boldsymbol{e}}$ (this is the case when corresponding $\hat{P}_{\boldsymbol{e}}$ has an odd number of $\hat{y}$ instances),[†] $\hat{P}_{\boldsymbol{e}} \cong \upsilon_{\boldsymbol{e}}\Gamma_{\boldsymbol{e}}$ possesses the same gradient as arbitrary rank excitation $\hat{E} = \hat{\Gamma}_{\boldsymbol{e}} D(\{\hat{\gamma}_{2p}\hat{\gamma}_{2p+1}\})$. Essentially, the iQCC ranking procedure identifies Majorana strings arising from any possible excitation on $|\phi_0\rangle$, up to rank $r = N_e$ (see Section A.3 for a discussion of the excitations associated to $\hat{P}$'s identified to have non-zero energy gradient for the initial Hamiltonian). We also establish that

$$\hat{E} \cong \upsilon_{\boldsymbol{e}}^{-1}\hat{P}_{\boldsymbol{e}} D(\{\hat{z}_p\}), \tag{A.22}$$

where $D(\{\hat{z}_p\}) \cong D(\{\hat{\gamma}_{2p}\hat{\gamma}_{2p+1}\})$ is the qubit-image of the projective polynomial. Physically, $\hat{P}_{\boldsymbol{e}}$ performs action of excitation $\hat{E}$, up to 1) the killing conditions of $\hat{E}$ due to absence of projector $D(\{\hat{z}_p\})$,[‡] and 2) a potential difference in phase factor. By point 1, we mean $\hat{P}_e$ and $\hat{E}$ transform populations of the same modes in general configuration $|\phi_f\rangle$, however, while $\hat{E}^{a,b,\dots c}_{ij,\dots k}$ kills $|\phi_f\rangle$ if $|\phi_f\rangle$ possesses any occupied modes in $\{a,b\dots c\}$ or any virtual modes in $\{i,j,\dots k\}$, action of $\hat{P}_{\boldsymbol{e}}$ simply switches modes $\{a,b\dots,c,i,j,\dots,k\}$ from virtual-to-occupied and vice versa. To illustrate, consider double excitation $\hat{E}^{32}_{01} = \hat{a}_3^\dagger\hat{a}_2^\dagger\hat{a}_0\hat{a}_1$. For corresponding $\hat{P}_{\boldsymbol{e}}$, we have $\hat{P}_{\boldsymbol{e}} \cong -\hat{\gamma}_0\hat{\gamma}_2\hat{\gamma}_4\hat{\gamma}_6$, which gives $\hat{P}_{\boldsymbol{e}} = \hat{y}_0\hat{x}_1\hat{y}_2\hat{x}_3$ in the JW representation. If $|\phi_f\rangle$ is configuration $|n_0n_1n_2n_3\rangle = |1100\rangle$, one has

$$\hat{E}^{32}_{01}|1100\rangle = |0011\rangle \tag{A.23}$$

$$\hat{P}_{\boldsymbol{e}}|1100\rangle = |0011\rangle\,. \tag{A.24}$$

However, if $|\phi_f\rangle = |1010\rangle$, we see

$$\hat{E}^{32}_{01}|1010\rangle = 0 \tag{A.25}$$

$$\hat{P}_{\boldsymbol{e}}|1010\rangle = -|0101\rangle\,. \tag{A.26}$$

As a consequence of the lack of killing conditions in $\hat{P}_{\boldsymbol{e}}$, there is no guarantee that rotation

[†]The discussion in Section A.2 includes how one can, in principle, modify $\upsilon_{\boldsymbol{e}}\hat{\Gamma}_{\boldsymbol{e}}$ such that its qubit image has an odd number of $\hat{y}$ instances while retaining equivalent gradient magnitude to $\hat{E}$.

[‡]See Section A.4 for how $\hat{E}$ inherits its killing conditions from projective polynomial $D(\{\hat{\gamma}_{2p}\hat{\gamma}_{2p+1}\})$.

generated by $\hat{P}_{\boldsymbol{e}}$ will preserve the particle number N_e, e.g.,

$$e^{\tau \hat{E}^{32}_{01}/2} |1110\rangle = \left(\hat{1} + \frac{\tau}{2}\hat{E}^{32}_{01}\right) |1110\rangle \tag{A.27}$$

$$= |1100\rangle \tag{A.28}$$

$$e^{-i\tau \hat{P}_{\boldsymbol{e}}/2} |1110\rangle = \left(\cos\left(\frac{\tau}{2}\right)\hat{1} - i\sin\left(\frac{\tau}{2}\right)\hat{P}_{\boldsymbol{e}}\right) |1110\rangle \tag{A.29}$$

$$= \cos\left(\frac{\tau}{2}\right) |1110\rangle + i\sin\left(\frac{\tau}{2}\right) |0001\rangle . \tag{A.30}$$

However, since $\hat{P}_e \cong \upsilon_{\boldsymbol{e}}\hat{\Gamma}_{\boldsymbol{e}}$'s are selected for inclusion in the iQCC unitary based on the energy derivative with respect to particular reference $|\phi_{0,f}\rangle$, particle number violations are expectedly insignificant.

A.2 $\hat{Z}$ gauge

In Section 2.2, it is elucidated that a Pauli term $\hat{P}_\alpha = \theta_\alpha \hat{X}_\alpha \hat{Z}_\alpha$ has energy derivative fully determined by $\hat{X}_\alpha$, so long as $\hat{Z}_\alpha$ has odd overlapping support with $\hat{X}_\alpha$. The freedom in choosing such $\hat{Z}_\alpha$ is exploited in Chapters 3 and 4 to reduce the resource requirements of the iQCC procedure. Here, we describe what such $\hat{Z}_\alpha$ gauge freedom of the gradients corresponds to in the fermionic picture.

For iQCC energy derivative Eq.(A.21), we can consider substituting arbitrary $\upsilon_{\boldsymbol{e}}\hat{\Gamma}_{\boldsymbol{e}}$ for some

$$\tilde{\upsilon}_{\boldsymbol{e}}\tilde{\Gamma}_{\boldsymbol{e}} = \upsilon_{\boldsymbol{e}}\hat{\Gamma}_{\boldsymbol{e}} \prod_{p\in G} \hat{\gamma}_{2p}\hat{\gamma}_{2p+1}, \tag{A.31}$$

for arbitrary $G \subseteq \{0, 1, \ldots, 2M-1\}$. We obtain

$$\frac{\partial E}{\partial \tau_{\boldsymbol{e}}}\Big|_{\tau_{\boldsymbol{e}}=0} = \mathrm{Im}\left(\tilde{\upsilon}_{\boldsymbol{e}} \langle\phi_{0,f}| \hat{H}\hat{\Gamma}_{\boldsymbol{e}} \prod_{p\in G} \hat{\gamma}_{2p}\hat{\gamma}_{2p+1} |\phi_{0,f}\rangle\right). \tag{A.32}$$

Noting Eq.(A.11), the modified derivative becomes

$$\frac{\partial E}{\partial \tau_{\boldsymbol{e}}}\Big|_{\tau_{\boldsymbol{e}}=0} = \mathrm{Im}\left((-i)^{|G|}\tilde{\upsilon}_{\boldsymbol{e}} \langle\phi_{0,f}| \hat{H}\hat{\Gamma}_{\boldsymbol{e}} \prod_{p\in G} (2\hat{n}_p - 1) |\phi_{0,f}\rangle\right) \tag{A.33}$$

$$= \mathrm{Im}\left((-i)^{|G|} \prod_{p\in G} \lambda_p \tilde{\upsilon}_{\boldsymbol{e}} \langle\phi_{0,f}| \hat{H}\hat{\Gamma}_{\boldsymbol{e}} |\phi_{0,f}\rangle\right), \tag{A.34}$$

where $\lambda_p = \langle\phi_{0,f}| (2\hat{n}_p - 1) |\phi_{0,f}\rangle \in \{1, -1\}$. Hence, $\prod_{p\in G} \lambda_p$ can not change the magnitude of the energy derivative. So long as $(-i)^{|G|}\tilde{\upsilon}_{\boldsymbol{e}}\hat{\Gamma}_{\boldsymbol{e}}$ is a purely imaginary-valued quantity, we recover identical magnitude with Eq.(A.21), using modified Majorana string Eq.(A.31).

Recalling that $(2\hat{n}_p - 1) \cong \hat{z}_p$ in the JW representation,[§] we have hence connected the $\hat{Z}$ gradient freedom in iQCC generators to the appending of $\hat{\gamma}_{2p}\hat{\gamma}_{2p+1}$ terms to the corresponding Majorana string. While ensuring $(-i)^{|G|}\tilde{\upsilon}_{\boldsymbol{e}}\hat{\Gamma}_{\boldsymbol{e}}$ is purely imaginary is slightly more involved in the fermionic representation, it is guaranteed in the qubit picture by enforcing the constraint that $\hat{Z}_\alpha$ has odd overlapping support with $\hat{X}_\alpha$.

A.3 Degree of Majorana string generators

For the initial electronic Hamiltonian, it is well-known that only double excitations $\hat{E}^{ab}_{ij}$ have non-vanishing matrix elements of the form $\langle\phi_{0,f}|\,\hat{H}\hat{E}^{a,b,\ldots c}_{i,j,\ldots k}\,|\phi_{0,f}\rangle$.[6] This is because $\hat{H}$ is a two-body operator, and hence any triple or higher rank excitations of the reference can not be coupled to the initial reference via $\hat{H}$. The single excitations have zero gradient due to the Brillouin theorem,[6] which establishes that the Hartree-Fock reference can not be coupled to a singly excited configuration through $\hat{H}$. Hence, the first iteration of iQCC screening procedure will identify Pauli terms $\hat{P}_{\boldsymbol{e}} \cong \upsilon_{\boldsymbol{e}}\hat{\Gamma}_{\boldsymbol{e}}$ arising from non-zero gradient doubles $\hat{E}^{ab}_{ij}$.

However, after the first iQCC dressing is performed, there can be more variety in the degree of Majorana strings identified to have non-zero gradients. Noting that new terms in the transformed Hamiltonian are generated by commutator $[\hat{P}_{\boldsymbol{e}}, \hat{H}]$, we consider this operator quantity in the fermionic representation, $\upsilon_{\boldsymbol{e}}[\Gamma_{\boldsymbol{e}}, \hat{H}]$. Since the original Hamiltonian consists of one- and two-body operators, it is represented as a linear combination of Majorana strings strictly of degree 2 and 4. The commutation rules[76] for two general Majorana strings $\hat{\Gamma}_{\boldsymbol{s}}$, $\hat{\Gamma}_{\boldsymbol{t}}$ can be expressed as

$$[\hat{\Gamma}_{\boldsymbol{s}}, \hat{\Gamma}_{\boldsymbol{t}}] = \left((-1)^{|\boldsymbol{s}||\boldsymbol{t}|+\boldsymbol{s}\cdot\boldsymbol{t}} - 1\right)\hat{\Gamma}_{\boldsymbol{t}}\hat{\Gamma}_{\boldsymbol{s}}. \tag{A.35}$$

and hence for non-commutative instances of iQCC generator $\upsilon_{\boldsymbol{e}}\hat{\Gamma}_{\boldsymbol{e}}$ and Majorana string $\hat{\Gamma}^{(h)}_{\boldsymbol{s}}$ from the Hamiltonian (arising when $|\boldsymbol{e}||\boldsymbol{s}| + \boldsymbol{e}\cdot\boldsymbol{s}$ is odd), the transformed $\hat{H}$ possesses terms proportional to $\hat{\Gamma}^{(h)}_{\boldsymbol{s}}\hat{\Gamma}_{\boldsymbol{e}}$, which can be of varying degree. For even Majorana degree d_h of $\hat{\Gamma}^{(h)}_{\boldsymbol{s}}$ and even Majorana degree d_e of $\hat{\Gamma}_{\boldsymbol{e}}$, they must possess an odd number of shared indices to be non-commutative. With minimum odd overlap of one, the degree of $\hat{\Gamma}^{(h)}_{\boldsymbol{s}}\hat{\Gamma}_{\boldsymbol{e}}$ is $d_h + d_e - 2$. Hence, after the initial dressing with Majorana strings of degree $2r = 4$ (they arose from double excitations, $r = 2$), one can identify degree 6 Majorana strings in the transformed Hamiltonian, which resulted from such a $\hat{\Gamma}_{\boldsymbol{e}}$ having an overlap of one with a degree 4 string $\Gamma^{(h)}_{\boldsymbol{s}}$. Hence, iteratively dressing will inevitably lead to the identifying of Pauli terms $P_{\boldsymbol{e}} \cong \upsilon_{\boldsymbol{e}}\hat{\Gamma}_{\boldsymbol{e}}$ with correspondence to arbitrary rank excitations.

[§]The other fermion-to-qubit mappings such as Bravyi Kitaev or parity may yield a product of $\hat{z}_{p'}$'s for $(2\hat{n}_p - 1)$ rather than a single $\hat{z}_p$ due to their non-local attribute. However, this does not change any of the conclusions made.

A.4 Properties of projector $D(\{\hat{\gamma}_{2p}\hat{\gamma}_{2p+1}\})$

In Eq.(A.22), we have connected the iQCC generators to fermionic excitations. Essentially, the Pauli term iQCC generators encode arbitrary excitations $\hat{E}^{ab\ldots c}_{ij\ldots k}$, up to the absence of an idempotent, diagonal polynomial of terms $\{\hat{\gamma}_{2p}\hat{\gamma}_{2p+1}\}_{p\in\{a,b,\ldots,c,i,j,\ldots,k\}}$. Such $D(\{\hat{\gamma}_{2p}\hat{\gamma}_{2p+1}\})$ gives rise to the substantially different algebraic properties of $\hat{P}_{\boldsymbol{e}}$ and $\hat{E}^{ab\ldots c}_{ij\ldots k}$, hence it is worth elucidating some of the properties of $D(\{\hat{\gamma}_{2p}\hat{\gamma}_{2p+1}\})$ itself. In Theorem 1 presented at the end of this section, we provide a proof of the idempotency for general $D(\{\hat{\gamma}_{2p}\hat{\gamma}_{2p+1}\})$.

Given the unitarity of general Majorana strings, Eq.(A.13) reveals interesting structure that $\hat{E}^{ab\ldots c}_{ij\ldots k} = \hat{U}\hat{P}$, i.e., any excitation can be decomposed into a unitary $\hat{U} = \Gamma_{\boldsymbol{e}}$ right-multiplied by a projector. The projector $\hat{P} = D(\{\hat{\gamma}_{2p}\hat{\gamma}_{2p+1}\})$ projects to the subspace complementary to the kernel of $\hat{E}^{ab\ldots c}_{ij\ldots k}$, i.e., they possess identical kernels. Denoting the kernel of operator $\hat{O}$ as $\ker(\hat{O})$, the $\ker(\hat{E}^{ab\ldots c}_{ij\ldots k})$ consists of all configurations where any modes in $a, b, \ldots c$ are occupied and any modes in $i, j \ldots k$ are empty. The fact that $\ker(\hat{P}) = \ker(\hat{E}^{ab\ldots c}_{ij\ldots k})$ is straightforward to see when we write $\hat{P} = \hat{U}^\dagger \hat{E}^{ab\ldots c}_{ij\ldots k}$. Due to unitarity of $\hat{U}^\dagger$, it has an empty kernel, $\dim(\ker(\hat{U}^\dagger)) = 0$, hence $\ker(\hat{P}) = \ker(\hat{U}^\dagger \hat{E}^{ab\ldots c}_{ij\ldots k}) = \ker(\hat{E}^{ab\ldots c}_{ij\ldots k})$.

To provide better intuition how $\hat{U}\hat{P}$ performs the excitation $\hat{E}^{ab\ldots c}_{ij\ldots k}$, using definition Eq.(A.1), one has for $\hat{U} = \hat{\Gamma}_{\boldsymbol{e}}$:

$$\hat{\Gamma}_{\boldsymbol{e}} = \prod_{p\in\{i,j,\ldots k,a,b,\ldots c\}} \left(\hat{a}_p + \hat{a}_p^\dagger\right) \tag{A.36}$$

$$= \hat{a}_a\hat{a}_b\ldots\hat{a}_c\hat{a}_i\hat{a}_j\ldots\hat{a}_k + \ldots + \hat{E}^{ab\ldots c}_{ij\ldots k} + \ldots + \hat{a}_a^\dagger\hat{a}_b^\dagger\ldots\hat{a}_c^\dagger\hat{a}_i^\dagger\hat{a}_j^\dagger\ldots\hat{a}_k^\dagger, \tag{A.37}$$

essentially giving a linear combination of 2^r fermionic products, where each product contains either $\hat{a}_p^\dagger$ or $\hat{a}_p$ for $p \in \{i, j, \ldots k, a, b, \ldots c\}$. However, for a general state $|\psi\rangle$ acted on $\hat{P}$, $\hat{P}|\psi\rangle$ only contains electronic configurations with modes $\{a, b, \ldots c\}$ unoccupied and modes $\{i, j, \ldots k\}$ occupied. Thus, any products in Eq.(A.37) possessing $\hat{a}_p$ for any $p \in \{a, b, \ldots, c\}$ or $\hat{a}_p^\dagger$ for any $p \in \{i, j, \ldots, k\}$, must kill $\hat{P}|\psi\rangle$. The only product in Eq.(A.37) which does not kill $\hat{P}|\psi\rangle$ is precisely $\hat{E}^{a,b,\ldots c}_{i,j,\ldots k}$. Indeed, this leads to $\hat{U}\hat{P}\hat{U}\hat{P} = 0$, as expected for $(\hat{E}^{ab\ldots c}_{ij\ldots k})^2 = 0$, since $\hat{U}\hat{P}|\psi\rangle \in \ker(\hat{P})$.

Theorem 1. *If* $D(\{\hat{\gamma}_{2p}\hat{\gamma}_{2p+1}\}) \equiv (-1)^r\hat{\Gamma}_{\boldsymbol{e}}\hat{E}^{ab\ldots c}_{ij\ldots k}$, *where*

$$\hat{E}^{ab\ldots c}_{ij\ldots k} = \hat{a}_a^\dagger\hat{a}_b^\dagger\ldots\hat{a}_c^\dagger\hat{a}_i\hat{a}_j\ldots\hat{a}_k \tag{A.38}$$

is a general rank r *excitation, and*

$$\hat{\Gamma}_{\boldsymbol{e}} = \hat{\gamma}_{2a}\hat{\gamma}_{2b}\ldots\hat{\gamma}_{2c}\hat{\gamma}_{2i}\hat{\gamma}_{2j}\ldots\hat{\gamma}_{2k}, \tag{A.39}$$

then $D(\{\hat{\gamma}_{2p}\hat{\gamma}_{2p+1}\})^2 = D(\{\hat{\gamma}_{2p}\hat{\gamma}_{2p+1}\})$.

Proof. In $D(\{\hat{\gamma}_{2p}\hat{\gamma}_{2p+1}\})^2 = \hat{\Gamma}_{\boldsymbol{e}}\hat{E}^{ab\ldots c}_{ij\ldots k}\hat{\Gamma}_{\boldsymbol{e}}\hat{E}^{ab\ldots c}_{ij\ldots k}$ we utilize explicit forms for the first instance

of $\hat{E}^{ab\ldots c}_{ij\ldots k}$ and the second instance of $\hat{\Gamma}_{\boldsymbol{e}}$:

$$D(\{\hat{\gamma}_{2p}\hat{\gamma}_{2p+1}\})^2 = \hat{\Gamma}_{\boldsymbol{e}}\hat{a}^\dagger_a\hat{a}^\dagger_b\ldots\hat{a}^\dagger_c\hat{a}_i\hat{a}_j\ldots\hat{a}_k\hat{\gamma}_{2a}\hat{\gamma}_{2b}\ldots\hat{\gamma}_{2c}\hat{\gamma}_{2i}\hat{\gamma}_{2j}\ldots\hat{\gamma}_{2k}\hat{E}^{ab\ldots c}_{ij\ldots k}. \tag{A.40}$$

Firstly, we can use the following useful anticommutation relations:

$$\{\hat{\gamma}_{2p},\hat{a}_q\} = \{\hat{\gamma}_{2p},\hat{a}^\dagger_q\} = \delta_{pq}. \tag{A.41}$$

Eq.(A.41) directly follows from Eq.(A.3) and that $\hat{a}_q = \frac{1}{2}(\hat{\gamma}_{2q} + i\hat{\gamma}_{2q+1})$ and $\hat{a}^\dagger_q = \frac{1}{2}(\hat{\gamma}_{2q} - i\hat{\gamma}_{2q+1})$. We use such anticommutation relations to sequentially transpose the Majorana operators in the second instance of $\hat{\Gamma}_{\boldsymbol{e}}$ such that each $\hat{\gamma}_{2p}$ is directly right-multiplying $\hat{a}^\dagger_p$ (or $\hat{a}_p$) in the first instance of $\hat{E}^{ab\ldots c}_{ij\ldots k}$. Since we are never transposing $\hat{\gamma}_{2p}$'s with $\hat{a}^\dagger_p$ or $\hat{a}_p$ in such a process, we use $\hat{a}_q\hat{\gamma}_{2p} = -\hat{\gamma}_{2p}\hat{a}_q$ and $\hat{a}^\dagger_q\hat{\gamma}_{2p} = -\hat{\gamma}_{2p}\hat{a}^\dagger_q$ for $p \neq q$. The total number of transpositions required is $r(2r-1)$, and we obtain

$$D(\{\hat{\gamma}_{2p}\hat{\gamma}_{2p+1}\})^2 = (-1)^{r(2r-1)}\hat{\Gamma}_{\boldsymbol{e}}\hat{a}^\dagger_a\hat{\gamma}_{2a}\hat{a}^\dagger_b\hat{\gamma}_{2b}\ldots\hat{a}^\dagger_c\hat{\gamma}_{2c}\hat{a}_i\hat{\gamma}_{2i}\hat{a}_j\hat{\gamma}_{2j}\ldots\hat{a}_k\hat{\gamma}_{2k}\hat{E}^{ab\ldots c}_{ij\ldots k}. \tag{A.42}$$

Finally, for each $\hat{a}_p\hat{\gamma}_{2p}$ and $\hat{a}^\dagger_p\hat{\gamma}_{2p}$, we again use Eq.(A.41), where $\hat{a}_p\hat{\gamma}_{2p} = 1 - \hat{\gamma}_{2p}\hat{a}_p$ and $\hat{a}^\dagger_p\hat{\gamma}_{2p} = 1 - \hat{\gamma}_{2p}\hat{a}^\dagger_p$, giving

$$D(\{\hat{\gamma}_{2p}\hat{\gamma}_{2p+1}\})^2 = (-1)^{r(2r-1)}\hat{\Gamma}_{\boldsymbol{e}}\left(\prod_{p\in\{a,b,\ldots,c\}} 1-\hat{\gamma}_{2p}\hat{a}^\dagger_p\right)\left(\prod_{p\in\{i,j,\ldots,k\}} 1-\hat{\gamma}_{2p}\hat{a}_p\right)\hat{E}^{ab\ldots c}_{ij\ldots k}. \tag{A.43}$$

Considering the expansion of the two products in Eq.(A.43), all non-identity terms vanish upon multiplication of $\hat{E}^{ab\ldots c}_{ij\ldots k}$ since one either has $\hat{a}^\dagger_p\hat{E}^{ab\ldots c}_{ij\ldots k}$ with $p \in \{a,b,\ldots,c\}$ or $\hat{a}_p\hat{E}^{ab\ldots c}_{ij\ldots k}$ with $p \in \{i,j,\ldots,k\}$, and both cases are zero due to the nilpotency of the fermionic creation/annihilation operators. Hence we obtain

$$D(\{\hat{\gamma}_{2p}\hat{\gamma}_{2p+1}\})^2 = (-1)^{r(2r-1)}\hat{\Gamma}_{\boldsymbol{e}}\hat{E}^{ab\ldots c}_{ij\ldots k}, \tag{A.44}$$

and since $r(2r-1)$ is odd (even) when r is odd (even), we obtain

$$D(\{\hat{\gamma}_{2p}\hat{\gamma}_{2p+1}\})^2 = (-1)^r\hat{\Gamma}_{\boldsymbol{e}}\hat{E}^{ab\ldots c}_{ij\ldots k} = D(\{\hat{\gamma}_{2p}\hat{\gamma}_{2p+1}\}). \tag{A.45}$$

□

Appendix B

Generating mutually anti-commuting sets of Pauli terms

In the standard iQCC generator selection procedure, one obtains the $\hat{x}$ strings $\vec{\mu}^{(i)}$ of the DIS (see Eq.(2.22)), from which Pauli terms with $\hat{X}_\alpha$ part given by a $\vec{\mu}^{(i)}$, typically taking one generator from each of the N_g highest gradient $\vec{\mu}^{(i)}$'s. To ensure the generators are purely imaginary, the only constraint on the $\hat{Z}_\alpha$ part is that it has odd overlapping support with $\hat{X}_\alpha$. For the ILC transformation, the constraint of mutual anti-commutativity poses additional difficulty in the generator selection process. A brute force search through all $\hat{Z}_\alpha$ components is unfeasible due to their exponential number, likewise heuristic graph techniques[16] are unfavorable due to the exponential number of vertices. Instead, we make use of the binary vector representation of Pauli terms and formulate the problem of finding N_g anti-commuting Pauli terms from the DIS partitions of interest as a system of linear equations.

In the binary vector representation, any N_q-qubit Pauli term $\hat{P}_k$ may be represented by two binary vectors of dimension N_q, $\vec{\nu}^{\,(k)}$ and $\vec{\mu}^{\,(k)}$, where their l^{th} components specify the single-qubit Pauli operator acting on the l^{th} qubit in $\hat{P}_k$, [67]

$$(\mu_l^{(k)}, \nu_l^{(k)}) = \begin{cases} (0,1) & l^{\text{th}} \text{ qubit is } \hat{z} \\ (1,0) & l^{\text{th}} \text{ qubit is } \hat{x} \\ (1,1) & l^{\text{th}} \text{ qubit is } \hat{y} \\ (0,0) & l^{\text{th}} \text{ qubit is identity} \end{cases} . \tag{B.1}$$

Noting that all Pauli terms with identical $\vec{\mu}$ vectors have the same energy gradient, we can formulate the problem of finding mutually anti-commuting elements from the DIS as follows: given $\{\vec{\mu}^{\,(k)}\}_{k=1}^{N_g}$, find the set of $\vec{\nu}$ vectors $\{\vec{\nu}^{\,(k)}\}_{k=1}^{N}$ such that the set of Pauli terms $\{\hat{P}_k\}_{k=1}^{N_g}$ represented by $\{(\vec{\mu}^{\,(k)}, \vec{\nu}^{\,(k)})\}_{k=1}^{N}$ are 1) mutually anti-commutative and 2) possessing an odd number of $\hat{y}$ instances. The condition of mutual anti-commutativity produces $N_g(N_g-1)/2$

linear equations on the $N_g N_q$ free variables in $\{\vec{\nu}^{\,(k)}\}_{k=1}^{N}$ of the form

$$A_{jk} \equiv \sum_{l=1}^{N_q} \mu_l^{(j)} \nu_l^{(k)} + \nu_l^{(j)} \mu_l^{(k)} \mod 2 = 1, \tag{B.2}$$

where $1 \leq j < k \leq N_g$. Note that Eq. (B.2) is equivalent to the $2N_q$-dimensional concatenated vectors $(\vec{\mu}^{\,(j)}\ \vec{\nu}^{\,(j)})$ and $(\vec{\mu}^{\,(k)}\ \vec{\nu}^{\,(k)})$ being non-orthogonal with respect to the symplectic inner product over the binary field, which implies corresponding $\hat{P}_j$ and $\hat{P}_k$ anti-commute. [67] The condition that all $\{\hat{P}_k\}_{k=1}^{N_g}$ have an odd number of $\hat{y}$ instances is easily enforced by demanding that each $(\vec{\mu}^{\,(k)}, \vec{\nu}^{\,(k)})$ pair has an odd number of $\mu_l^{(k)} = \nu_l^{(k)} = 1$ occurrences. As N_g linear equations of the $\{\vec{\nu}^{\,(k)}\}_{k=1}^{N_g}$ variables,

$$O_k \equiv \sum_{l=1}^{N_q} \mu_l^{(k)} \nu_l^{(k)} \mod 2 = 1, \tag{B.3}$$

where $1 \leq k \leq N_g$. Collectively, $\{A_{jk}\}_{j<k}^{N}$ and $\{O_k\}_{k=1}^{N}$ are $N_g(N_g+1)/2$ binary linear equations, for which finding the $N_g N_q$ free variables in $\{\vec{\nu}^{\,(k)}\}_{k=1}^{N_g}$ which satisfy their system yields $\{(\vec{\mu}^{\,(k)}, \vec{\nu}^{\,(k)})\}_{k=1}^{N_g} \rightarrow \{\hat{P}_k\}_{k=1}^{N_g}$ which are mutually anti-commuting elements of the corresponding input DIS partitions governed by $\{\vec{\mu}^{\,(k)}\}_{k=1}^{N_g}$ with gradients $\{g_k\}_{k=1}^{N_g}$.

The linear system of equations can be formulated in matrix notation, for which standard implementations of binary gaussian elimination can be readily employed for finding a solution, if one exists. For instance, let $\boldsymbol{z}$ denote the $(N_g N_q)$-dimensional concatenated $\vec{\nu}$ vectors, $\boldsymbol{z} = \left(\vec{\nu}^{\,(1)}\ \ldots\ \vec{\nu}^{\,(N_g)}\right)^T$, and let $\mathbf{1}$ denote the $[N_g(N_g+1)/2]$-dimensional vector with unit value in all components. Then the system of linear equations can be written

$$\boldsymbol{A}\boldsymbol{z} = \mathbf{1}, \tag{B.4}$$

where $\boldsymbol{A}$ is a binary matrix of $N_g N_q$ columns and $N_g(N_g+1)/2$ rows. Each of the $N_g(N_g-1)/2$ anti-commutative constraints A_{jk}'s are encoded in a row of $\boldsymbol{A}$. The row corresponding to constraint A_{jk} has the form $(\mathbf{0}\ldots\vec{\mu}^{(k)}\ldots\mathbf{0}\ldots\vec{\mu}^{(j)}\ldots\mathbf{0})$ where $\mathbf{0}$ is the N_q-dimensional zero vector, $\vec{\mu}^{(k)}$ and $\vec{\mu}^{(j)}$ indices range from $[(j-1)N_q+1]$ to jN_q and $[(k-1)N_q+1]$ to kN_q respectively. The remaining $\hat{y}$ parity constraints are encoded in the remaining N_g rows of $\boldsymbol{A}$, where row corresponding to constraint O_k takes the form $(\mathbf{0}\ldots\vec{\mu}^{\,(k)}\ldots\mathbf{0})$, with $\vec{\mu}^{\,(k)}$ from column indices $[(k-1)N_q+1]$ to kN_q.

Finite field Gaussian elimination requires $O(m^3)$ binary arithmetic operations for m variables, [77] hence obtaining solution to Eq. (B.4) can be done in $O(N_g^3 N_q^3)$ time. Since N_g is bound by $O(N_q)$,[66] the worst-case scaling of this procedure is $O(N_q^6)$. There exist instances of $\boldsymbol{A}$ for which $\mathbf{1}$ is not in its image, i.e. Eq. (B.4) has no solution. In cases where no solution exists, we employ greedy method of replacing the lowest gradient $\vec{\mu}^{(k)}$ considered with the highest gradient unconsidered $\vec{\mu}^{(k')}$ until a solution to the corresponding $\boldsymbol{A}$ can be

found. If there exists no unconsidered $\vec{\mu}^{(k')}$'s during this greedy process greedy process, the lowest is simply removed, and an effective $N'_g = N_g - 1$ is used.

Appendix C

Optimization of the iQCC-ILC unitary

The state generated by acting a iQCC-ILC unitary, $\hat{U}_{\mathrm{ILC}}$, on the reference computational basis state (Slater determinant) $|\phi_0\rangle$ may be written

$$\hat{U}_{\mathrm{ILC}} |\phi_0\rangle = \exp\left(-i\tau \sum_{j=1}^{N_g} d_j \hat{P}_j\right) |\phi_0\rangle \tag{C.1}$$

$$\equiv c_0 |\phi_0\rangle - i \sum_{j=1}^{N_g} c_j \hat{P}_j |\phi_0\rangle , \tag{C.2}$$

where $c_0 = \cos(\tau)$ and $c_j = \sin(\tau) d_j$ for $j > 0$. Letting $|\Phi_0\rangle \equiv |\phi_0\rangle$, and $|\Phi_j\rangle \equiv \hat{P}_j |\phi_0\rangle$ for $j > 0$, one has by construction,

$$\langle \Phi_i | \Phi_j \rangle = \delta_{ij}, \tag{C.3}$$

since all $\hat{P}_j$ in the involutory combination are necessarily characterized by different flip indices (they originate from different DIS partitions and hence have distinct $\hat{X}_j$ components) and $|\phi_0\rangle$ is a computational basis state. Finding the optimal amplitudes and energy minimum can hence be done by generating Hamiltonian matrix $\boldsymbol{H}$ in the set of N_g+1 orthogonal basis functions $\{|\Phi_j\rangle\}_{j=0}^{N_g}$,

$$H_{ij} \equiv \langle \Phi_i | \hat{H} | \Phi_j \rangle , \tag{C.4}$$

from which the energy minimum and optimal parameters $\{c_j\}_{j=0}^{N_g}$ can be obtained via standard orthogonal eigen-problem formulation. The imaginary phase of the summed over configurations in Eq. (C.2) can be explicitly accounted for by modifying the Hamiltonian

matrix by element-wise multiplication, $\bar{\boldsymbol{H}} = \boldsymbol{M} \circ \boldsymbol{H}$, where

$$\boldsymbol{M} = \begin{pmatrix} 1 & -\boldsymbol{i}_{1\times N} \\ \boldsymbol{i}_{N\times 1} & \boldsymbol{1}_{N\times N} \end{pmatrix}. \tag{C.5}$$

The optimal coefficients $\{c_j\}_{i=0}^{N_g}$ in Eq. (C.2) may then be obtained by solving for the ground state solution of

$$\bar{\boldsymbol{H}}\boldsymbol{c} = E\boldsymbol{c}, \tag{C.6}$$

where E corresponds to the energy minimum for the iQCC-ILC unitary,

$$E = \min_{\tau,\boldsymbol{d}} \langle\phi_0| \hat{U}_{\mathrm{ILC}}^{\dagger}(\tau,\boldsymbol{d})\hat{H}\hat{U}_{\mathrm{ILC}}(\tau,\boldsymbol{d}) |\phi_0\rangle . \tag{C.7}$$

Unique matrix elements of $\boldsymbol{H}$ may be efficiently obtained on a classical computer due to the polynomial scaling of evaluating computational basis state (or more generally, QMF) expectation values as trigonometric polynomial functions of the $2N_q$ Bloch/Euler angles. [78] Once Eq. (C.6) has been solved, the iQCC-ILC parameters $\{\tau, d_1, \ldots d_N\}$ may be extracted from ground eigenvector $\boldsymbol{c}$ by obtaining amplitude τ as

$$\tau = \arccos(c_0), \tag{C.8}$$

which can then be used to obtain $\{d_j\}_{j=1}^{N_g}$,

$$d_j = \frac{c_j}{\sin\tau}. \tag{C.9}$$

From Eq. (C.9), $\{d_j\}_{j=1}^{N_g}$ are undefined when $\tau = 0$ (modulo π), however, such a scenario does not happen in practice. This is a result of all Pauli terms being selected from the DIS, and hence have non-zero energy gradient evaluated at $\tau = 0$, ensuring the optimized τ amplitude will be non-zero.

The procedure described so far will obtain the optimal ILC amplitudes for the initial computational basis state reference. Simultaneous relaxation of the reference Bloch angles and ILC amplitudes can be accomplished on a classical computer by an multiconfigurational self consistent field (MCSCF)-like two step iterative procedure: 1) solve generalized eigenproblem for subspace Hamiltonian matrix resolved in the N_g+1 states to obtain the current-step optimal ILC amplitudes $\{\tau, d_1, \ldots, d_{N_g}\}$, and 2) dress initial Hamiltonian $\hat{H}$ with the iQCC-ILC unitary using obtained amplitudes of the current step

$$\tilde{H} = \hat{U}_{\mathrm{ILC}}^{\dagger}(\tau,\boldsymbol{d})\hat{H}\hat{U}_{\mathrm{ILC}}(\tau,\boldsymbol{d}), \tag{C.10}$$

then perform Bloch angle optimization with respect to $\tilde{H}$. The current-step optimized

mean-field state is then used to update the subspace basis of Step 1. The two steps are then repeated until convergence of the energy, yielding the relaxed N_g ILC amplitudes and $2N_q$ Bloch angles. Over the course of optimization, the reference states may violate z-collinearity, resulting in subspace basis states $|\Phi_j\rangle$'s being no longer being generally orthogonal. If relaxation of QMF is considered, the next iteration of Step 1 must hence be accomplished by solving non-orthogonal generalization of Eq. (C.6),

$$\bar{\boldsymbol{H}}\boldsymbol{c} = E\bar{\boldsymbol{S}}\boldsymbol{c}, \tag{C.11}$$

where $\bar{\boldsymbol{S}} = \boldsymbol{M} \circ \boldsymbol{S}$ and $\boldsymbol{S}_{ij} = \langle \Phi_i | \Phi_j \rangle$.

Appendix D

Approximate probabilistic algorithm for finding growth minimizing Pauli terms

Algorithm 2 Probabilistic variant of Algorithm 1.

Input $\mathcal{H}$, $\vec{\mu}^{(k)}$, r, N_{samples}

Output smallest growth element $\hat{P}_\alpha$ with gradient g_α and growth $\gamma^{(k)}$

obtain set Ω
initiate $\mathcal{C} = \{\}$
initiate `valid_pairs` $= \{\}$
for $\vec{\mu}^{(i)}, \vec{\mu}^{(j)} \in \Omega$, $i < j$ **do**
    if $\vec{\mu}^{(i)} + \vec{\mu}^{(j)} \mod 2 = \vec{\mu}^{(k)}$ **then**
        update `valid_pairs` $\rightarrow$ `valid_pairs` $\cup \{(\vec{\mu}^{(i)}, \vec{\mu}^{(j)})\}$
    end if
end for
for sample in $1, 2, \ldots N_{\text{samples}}$ **do**
    uniformly select $(\vec{\mu}^{(i)}, \vec{\mu}^{(j)}) \in$ `valid_pairs`
    uniformly select $\hat{Z}_l^{(i)} \in \mathcal{Z}_i$, $\hat{Z}_{l'}^{(j)} \in \mathcal{Z}_j$
    compute $\hat{C} = [\hat{Z}_l^{(i)} \hat{X}_i, \hat{Z}_{l'}^{(j)} \hat{X}_j]$
    if $\hat{C} \neq 0$ **then**
        update $\mathcal{C} \rightarrow \mathcal{C} + \{\hat{C}\}$
    end if
end for
count occurrences $m_{\mathcal{C}}(\hat{P}_\alpha)$ for each unique $\hat{P}_\alpha \in \mathcal{C}$
for r highest occurring $\hat{P}_\alpha \in \mathcal{C}$ **do**
    compute $|\mathcal{H}_A^{(\alpha)}|$
    compute and save $\gamma_\alpha = \left| [\hat{P}_\alpha, \mathcal{H}] / \mathcal{H}_A^{(\alpha)} \right|$
end for
select $\hat{P}_{\alpha'}$ with smallest $\gamma_{\alpha'}$. Then $\gamma^{(k)} = \gamma_{\alpha'}$

Table D.1: A comparison of performance between the deterministic ("Det.") and probabilistic ("Prob.") minimum growth search algorithms of Algorithm 1 and Algorithm 2 respectively. For the probabilistic algorithm, $N_{\text{samples}} = M$ was used. For all examples tabulated, $\gamma^{(k)}$ was obtained for the highest gradient. Observed $\gamma^{(k)}$ values for the probabilistic algorithm are the averages of the lowest growth values computed over ten runs, with standard deviations in brackets. For each effective Hamiltonian, all ten runs find the minimal growth element found in the deterministic algorithm, hence all standard deviations are zero.

Iteration	M	Algorithm	N_{query}	Observed $\gamma^{(k)}$
0	247	Det.	540	112
		Prob.	247	112.0(0.0)
5	1.24×10^3	Det.	1.37×10^4	130
		Prob.	1.24×10^3	130.0(0.0)
10	4.47×10^3	Det.	1.29×10^5	878
		Prob.	4.47×10^3	878.0(0.0)
20	2.10×10^4	Det.	1.99×10^6	2263
		Prob.	2.10×10^4	2263.0(0.0)

For problem instances featuring large number of Hamiltonian terms M, Algorithm 1 can be bottlenecked when computing $\mathcal{C}_{ij}$ for all valid x-string pairs. To alleviate time and memory requirements of this step, we propose a modified version of Algorithm 1 which utilizes uniform sampling. First, we uniformly sample valid x-string pairs, i.e. uniformly select which set we sample from in $\{\mathcal{C}_{ij}\}_{(ij)}$. Then, instead of computing the selected $\mathcal{C}_{ij}$ set exhaustively, we sample an element of $\mathcal{C}_{ij}$ by uniformly sampling l and $l^{'}$ (see Eq. (4.13)). The resulting commutator constitutes a "sample" and is saved if it is non-zero. Since generally growths are negatively correlated with multiplicities in $\bigcup_{(ij)} \mathcal{C}_{ij}$, and higher multiplicity elements are uniformly sampled from $\bigcup_{(ij)} \mathcal{C}_{ij}$ with higher probability (the probability of sampling a Pauli product $\hat{P}_\alpha$ is given by $\sum_{(ij)} m_{\mathcal{C}_{ij}}(\hat{P}_\alpha)/|\sum_{(ij)} \mathcal{C}_{ij}|$), we find that this probabilistic variant of Algorithm 1 finds the lowest growth element with high success rate and with dramatically less resources. We found empirically that setting the number of samples to be a linear function of M successfully finds the lowest growth element for systems studied in this work. Similarly to Algorithm 1, we then compute $|\mathcal{H}_A^{(k)}|$ for the r most sampled Pauli products. However, we have not computed all of $\mathcal{C}$, rather we have only sampled elements from it N_{samples} times. Hence, the values of $m_{\mathcal{C}}(\hat{P}_\alpha)$ are not readily available. Thus we compute γ_α as $\left|[\hat{P}_\alpha, \mathcal{H}]/\mathcal{H}_A^{(\alpha)}\right|$. We summarize the probabilistic routine in Algorithm 2.

A comparison between the deterministic and probabilistic algorithms is summarized in Table D.1. Both algorithms were applied to the CAS(6e, 6o) N_2 system at bond length $R =$ 1.5 Å in the cc-pVDZ basis. Effective Hamiltonians at various iterations of the $s(1, 0)$ iQCC procedure applied to this system were used for assessment. The number of Hamiltonian Pauli products with coefficient values with magnitude above 10^{-8} is denoted by M. The

main cost of the algorithms is encapsulated by N_{query}, which refers to the number of times the algorithm must query the commutator function between two Pauli products. Note that non-zero commutators are saved, hence time and memory requirements of both algorithms scale with N_{query}. For both deterministic and probabilistic algorithms, $r = \lceil \log_2 M \rceil$ was used. We observe that even for large scale effective Hamiltonians, setting $N_{\text{samples}} = M$ in Algorithm 2 always finds the lowest growth element found in the deterministic approach of 1. This is significant, since in this case we query the commutator function between two Pauli products $N_{\text{query}} = M$ times using Algorithm 2, versus $N_{\text{query}} = O(M^2)$ using Algorithm 1. In practice, we observe up to $\sim 100\times$ reductions in the number of commutator queries resulting in immensely reduced time and memory requirements, while maintaining identical performance in terms of the lowest observed growth element.

Appendix E

DIS construction for general mixed states

E.1 Multireference screening

Herein, we derive a generalization of the DIS construction to cases where the reference state is a linear combination of electronic configurations.

Let $\mathcal{P} = \{\phi_i\}_{i=1}^{L}$ denote a model space of L Slater determinants. We consider a multi-configurational state in the span of $\mathcal{P}$,

$$|I\rangle = \sum_{\phi_j \in \mathcal{P}} c_j^{(I)} |\phi_j\rangle , \tag{E.1}$$

where $c_j^{(I)}$ are real and normalized coefficients, as the reference state in the QCC procedure. We wish to identify Pauli terms which possess high absolute values of the energy gradient,

$$\frac{\partial E_I}{\partial \tau_\alpha}\Big|_0 = \frac{\partial}{\partial \tau_\alpha} \langle I| e^{i\tau_\alpha \hat{P}_\alpha/2} \hat{H} e^{-i\tau_\alpha \hat{P}_\alpha/2} |I\rangle \Big|_0 \tag{E.2}$$

$$= \mathrm{Im} \langle I| \hat{H}\hat{P}_\alpha |I\rangle . \tag{E.3}$$

We refer to the absolute value of Eq.(E.2) as $g_\alpha^{(I)}$. Expanding the multideterminantal reference state as Eq.(E.1), we obtain gradient expression

$$\frac{\partial E_I}{\partial \tau_\alpha}\Big|_0 = \sum_{j,k=1}^{L} c_j^{(I)} c_k^{(I)} \,\mathrm{Im} \langle \phi_j| \hat{H}\hat{P}_\alpha |\phi_k\rangle , \tag{E.4}$$

and can introduce the factorization of arbitrary $\hat{P}_\alpha$ as $\hat{P}_\alpha = \theta_\alpha \hat{X}_\alpha \hat{Z}_\alpha$ to obtain

$$\frac{\partial E_I}{\partial \tau_\alpha}\Big|_0 = \sum_{j,k=1}^{L} c_j^{(I)} c_k^{(I)} \,\mathrm{Im}\left(\theta_\alpha \langle\phi_j| \hat{H}\hat{X}_\alpha \hat{Z}_\alpha |\phi_k\rangle\right) \tag{E.5}$$

$$= \sum_{j,k=1}^{L} c_j^{(I)} c_k^{(I)} \lambda_k^{(\alpha)} \,\mathrm{Im}\left(\theta_\alpha \langle\phi_j| \hat{H}\hat{X}_\alpha |\phi_k\rangle\right), \tag{E.6}$$

where $\hat{Z}_\alpha |\phi_k\rangle = \lambda_k^{(\alpha)} |\phi_k\rangle$. Now, we note that the linear combination of matrix elements in Eq.(E.6) can be re-expressed as a sum of (non-Hermitian) expectation values, if we re-write e.g. the bra states as transformations of their corresponding ket states. The computational basis states are transformed to one another by products of Pauli-$\hat{x}$ operations, hence one can always find the coupling $\hat{x}$ product $\hat{X}_{jk}$ which satisfies $\hat{X}_{jk}|\phi_k\rangle = |\phi_j\rangle$ (note that $\hat{X}_{jk} = \hat{X}_{kj}$). We obtain

$$\frac{\partial E_I}{\partial \tau_\alpha}\Big|_0 = \sum_{j=1}^{L}\sum_{k=1}^{L} c_j^{(I)} c_k^{(I)} \lambda_k^{(\alpha)} \,\mathrm{Im}\left(\theta_\alpha \langle\phi_k| \hat{X}_{jk}\hat{H}\hat{X}_\alpha |\phi_k\rangle\right) \tag{E.7}$$

$$= \sum_{k=1}^{L} c_k^{(I)} \lambda_k^{(\alpha)} \,\mathrm{Im}\left(\theta_\alpha \langle\phi_k| \sum_{j=1}^{L} c_j^{(I)} \hat{X}_{jk}\hat{H}\hat{X}_\alpha |\phi_k\rangle\right) \tag{E.8}$$

$$= \mathrm{Im}(\theta_\alpha) \sum_{k=1}^{L} c_k^{(I)} \lambda_k^{(\alpha)} \langle\phi_k| \sum_{j=1}^{L} c_j^{(I)} \hat{X}_{jk}\hat{H}\hat{X}_\alpha |\phi_k\rangle \tag{E.9}$$

$$= \mathrm{Im}(\theta_\alpha) \sum_{k=1}^{L} c_k^{(I)} \lambda_k^{(\alpha)} \langle\phi_k| \hat{H}_k^{(I)} \hat{X}_\alpha |\phi_k\rangle, \tag{E.10}$$

where in the last equality we have defined $\hat{H}_k^{(I)} \equiv \sum_{j=1}^{L} c_j^{(I)} \hat{X}_{jk}\hat{H}$. In going from Eq.(E.8) to Eq.(E.9), we make the observation that left- and right-multiplication of real-valued $\hat{H}$ by a general product of $\hat{x}$ operators must conserve realness. Since our starting configurations are assumed to be real, the expectation values in Eq.(E.9) must be strictly real, and hence the only potentially complex quantity is θ_α.

For easing of notation, let

$$\Xi_k^{(\alpha)} \equiv c_k^{(I)} \langle\phi_k| \hat{H}_k^{(I)} \hat{X}_\alpha |\phi_k\rangle, \tag{E.11}$$

and one can write $g_\alpha^{(I)}$ as

$$g_\alpha^{(I)} = \left|\mathrm{Im}(\theta_\alpha) \sum_{k=1}^{L} \lambda_k^{(\alpha)} \Xi_k^{(\alpha)}\right|. \tag{E.12}$$

Importantly, we have arrived at an expression for $g_\alpha^{(I)}$ which consists of a sum of K terms,

where each term only explicitly depends on a single reference state $|\phi_k\rangle$. This lets us straightforwardly construct the DIS for multideterminantal state $|I\rangle$ by considering L analogues to the standard single-reference DIS of Sec 2.2 for the $\Xi_k^{(\alpha)}$'s. We call such analogues "component" DISs (CDISs), as they explicitly only depend on one component of $|I\rangle$, and yield a component of the total vector-valued DIS for $|I\rangle$. To obtain the CDIS for the component $|\phi_k\rangle$, one can cast Eq.(E.11) as,

$$\Xi_k^{(\alpha)} = c_k^{(I)} \langle\phi_k| \left(\sum_{j=1}^{L} c_j^{(I)} \hat{X}_{jk} \hat{H} \right) \hat{X}_\alpha |\phi_k\rangle \tag{E.13}$$

$$= c_k^{(I)} \langle\phi_k| \left[\sum_{j=1}^{L} c_j^{(I)} \hat{X}_{jk} \sum_i \left(\sum_l \eta_l^{(i)} \hat{Z}_l^{(i)} \right) \hat{X}_i \right] \hat{X}_\alpha |\phi_k\rangle \tag{E.14}$$

$$= c_k^{(I)} \sum_{j=1}^{L} \sum_i c_j^{(I)} \langle\phi_k| \hat{X}_{jk} \left(\sum_l \eta_l^{(i)} \hat{Z}_l^{(i)} \right) \hat{X}_i \hat{X}_\alpha |\phi_k\rangle , \tag{E.15}$$

and inspecting the expectation value $\langle\phi_k| \hat{X}_{jk} \left(\sum_l \eta_l^{(i)} \hat{Z}_l^{(i)} \right) \hat{X}_i \hat{X}_\alpha |\phi_k\rangle$, it is clear that it vanishes unless $\hat{X}_{jk} |\phi_k\rangle = \hat{X}_i \hat{X}_\alpha |\phi_k\rangle$, due to the orthogonality of computational basis states. To satisfy this, one needs $\hat{X}_{jk} = \hat{X}_i \hat{X}_\alpha \to \hat{X}_\alpha = \hat{X}_i \hat{X}_{jk}$. In essence, the candidate Pauli term's $\hat{x}$ component $\hat{X}_\alpha$ must be equal to at least one instance of $\hat{X}_i \hat{X}_{jk}$, otherwise $\Xi_k^{(\alpha)} = 0$. Further, letting $\vec{\mu}^{[jk]}$ denote the $\hat{x}$ string representing $\hat{X}_{jk}$, and letting $\vec{\mu}^{(i,[jk])} = \vec{\mu}^{[jk]} + \vec{\mu}^{(i)}$,

$$\Xi_k^{(\alpha)} = c_k^{(I)} \sum_{j=1}^{L} \sum_i c_j^{(I)} \langle\phi_k| \left(\sum_l \eta_l^{(i)} \hat{Z}_l^{(i)} \right) |\phi_k\rangle \, \boldsymbol{\delta}_{\vec{\mu}^{(i,[jk])}, \vec{\mu}^{(\alpha)}} \tag{E.16}$$

$$= c_k^{(I)} \sum_{j=1}^{L} \sum_i c_j^{(I)} \left(\sum_l \eta_l^{(i)} \lambda_{l,\phi_k}^{(i)} \right) \boldsymbol{\delta}_{\vec{\mu}^{(i,[jk])}, \vec{\mu}^{(\alpha)}} , \tag{E.17}$$

where $\lambda_{l,\phi_k}^{(i)} = \langle\phi_k| \hat{Z}_l^{(i)} |\phi_k\rangle \in \{1, -1\}$. Hence, for any $\hat{X}_i \hat{X}_{jk} = \hat{X}_\alpha$ (i.e., $\boldsymbol{\delta}_{\vec{\mu}^{(i,[jk])}, \vec{\mu}^{(\alpha)}} = 1$), $\Xi_k^{(\alpha)}$ receives contribution $c_k^{(I)} c_j^{(I)} \sum_l \eta_l^{(i)} \lambda_{l,\phi_k}^{(i)}$. The CDIS for $\Xi_k^{(\alpha)}$ is constructed as follows. Firstly, obtain the set of x string vectors $\vec{\mu}^{(i)}$ corresponding to unique $\hat{x}$ tensor products found in $\hat{X}_{jk} \hat{H}$, for all $j \leq k$. We refer to the set of unique $\vec{\mu}^{(i)}$ found in $\hat{X}_{jk} \hat{H}$ as Ω_{jk}, and $\Omega_{jk} = \Omega_{kj}$. Referring to Ω as the set of $\hat{x}$ strings for $\hat{H}$, Ω_{jk} can be directly obtained from Ω by simply adding $\vec{\mu}^{[jk]}$ element-wise to Ω:

$$\Omega_{jk} = \{\vec{\mu}^{(i)} + \vec{\mu}^{[jk]} : \vec{\mu}^{(i)} \in \Omega\}. \tag{E.18}$$

We refer to the set of unique $\hat{x}$ strings found across all $\hat{X}_{jk}\hat{H}$ instances as

$$\Omega_I = \bigcup_{j \leq k} \Omega_{jk}. \tag{E.19}$$

For every Ω_{jk} in Eq.(E.19) , compute the contribution to $\Xi_k^{(\alpha)}$ for every $\vec{\mu}^{(i,[jk])} \in \Omega_{jk}$ as $c_k^{(I)} c_j^{(I)} \sum_l \eta_l^{(i)} \lambda_{l,\phi_k}^{(i)}$. This yields us the CDIS for ϕ_k,

$$\mathcal{C}_k = \left\{ \left(\vec{\mu}^{(\alpha)}, \Xi_k^{(\alpha)} \right) : \vec{\mu}^{(\alpha)} \in \Omega_I \right\}. \tag{E.20}$$

The DIS for multideterminantal state $|I\rangle$ is then written

$$\mathcal{D}^{(I)} = \left\{ \left(\vec{\mu}^{(\alpha)}, \vec{\boldsymbol{\Xi}}^{(\alpha)} \right) : \vec{\mu}^{(\alpha)} \in \Omega_I \right\}, \tag{E.21}$$

where $\vec{\boldsymbol{\Xi}}^{(\alpha)} = \left\{ \Xi_0^{(\alpha)}, \Xi_1^{(\alpha)}, \ldots, \Xi_{L-1}^{(\alpha)} \right\}$. That is, each candidate $\hat{x}$ string $\vec{\mu}^{(\alpha)}$ in Ω_I is assigned a L-dimensional vector quantity, $\vec{\boldsymbol{\Xi}}^{(\alpha)}$. Note that Eq.(E.21) is exactly equivalent to Eq.(2.22) when $L = 1$, and hence the developments here are a generalization of the standard single-reference iQCC (SR-iQCC) generator screening procedure. However, there are important distinctions from the standard procedure when $L > 1$. Mainly, when $L = 1$ (as in the SR-iQCC procedure), the DIS directly informs us the exact gradient magnitude offered by an X_α so long as Z_α is chosen to produce an odd number of $\hat{y}$ instances, an easily satisfiable constraint. For $L > 1$, the connection between $\vec{\boldsymbol{\Xi}}^{(\alpha)}$ and g_α is less straightforward. This is because of the role of relative phases $\{\lambda_k^{(\alpha)}\}_{k=0}^{L-1}$ in Eq.(E.12), which are determined by choice of $\hat{Z}_\alpha$. If the sign of all terms in Eq.(E.12) can be aligned, and $\hat{Z}_\alpha$ also leads to an odd number of $\hat{y}$ instances in $\hat{P}_\alpha$, then one has $g_\alpha = \sum_{k=0}^{L-1} \left| \Xi_k^{(\alpha)} \right|$. However, such a case is not always possible, i.e., there may not exist a $\hat{Z}_\alpha$ which can make the signs of all $\{\lambda_k^{(\alpha)} \Xi_k^{(\alpha)}\}_{k=0}^{L-1}$ uniform, while also satisfying the odd-$\hat{y}$ parity constraint. Thus, $\mathcal{D}^{(I)}$ informs us of the X_α parts which will potentially lead to non-zero g_α for multideterminantal reference $|I\rangle$, however, given $\hat{X}_\alpha$, finding the optimal $\hat{Z}_\alpha$ to produce the highest possible g_α, is far less trivial than in the SR-iQCC screening. In Appendix F, we outline the optimal and efficient strategies for finding $\hat{Z}_\alpha$ to yield generator $\hat{P}_\alpha$ with high g_α employed in this work.

E.2 Extending to state-averaged ensembles

In the previous section of this Appendix, we derived the procedure for finding the DIS for multiconfigurational reference state $|I\rangle$, which in conjunction with a phase alignment procedure from Appendix F, yields iQCC generators $\hat{P}_\alpha$ of high energy gradient g_α. Here, we show a simple extension of the MR-iQCC screening procedure to obtain the DIS (i.e., the set of possible $\hat{P}_\alpha$ $\hat{x}$-strings $\vec{\mu}^{(\alpha)}$ which potentially lead to non-zero g_α) for a state-averaged mixture of generally multiconfigurational references. We refer to this procedure as the

MRSA-iQCC screening. This procedure also applies to when the state-averaged mixture is over a set of single electronic configurations. It also reduces to the MR-iQCC screening procedure described in the last section when $N_s = 1$. It is hence the most generalized generator screening procedure derived in this work.

The gradient of state-averaged energy Eq.(5.3) is

$$\frac{\partial E_{\mathrm{SA}}}{\partial \tau_\alpha}\Big|_0 = \sum_{I=1}^{N_s} w_I \frac{\partial E_I}{\partial \tau_\alpha}\Big|_0. \tag{E.22}$$

Inserting Eq.(E.10) for each MR state gradient, one obtains

$$\frac{\partial E_{\mathrm{SA}}}{\partial \tau_\alpha}\Big|_0 = \mathrm{Im}(\theta_\alpha) \sum_{I=1}^{N_s} \sum_{k=1}^{L} w_I c_k^{(I)} \lambda_k^{(\alpha)} \langle \phi_k | \hat{H}_k^{(I)} \hat{X}_\alpha | \phi_k \rangle \tag{E.23}$$

$$= \mathrm{Im}(\theta_\alpha) \sum_{k=1}^{L} \lambda_k^{(\alpha)} \langle \phi_k | \sum_{I=1}^{N_s} w_I c_k^{(I)} \hat{H}_k^{(I)} \hat{X}_\alpha | \phi_k \rangle . \tag{E.24}$$

Upon further inspection, recalling $\hat{H}_k^{(I)} \equiv \sum_{j=1}^{L} c_j^{(I)} \hat{X}_{jk} \hat{H}$, we see we have only introduced a re-weighting of $\hat{X}_{jk}\hat{H}$'s going from the state-specific to state-average gradient,

$$\frac{\partial E_{\mathrm{SA}}}{\partial \tau_\alpha}\Big|_0 = \mathrm{Im}(\theta_\alpha) \sum_{k=1}^{L} \lambda_k^{(\alpha)} \langle \phi_k | \sum_{I=1}^{N_s} \sum_{j=1}^{L} w_I c_k^{(I)} c_j^{(I)} \hat{X}_{jk} \hat{H} \hat{X}_\alpha | \phi_k \rangle \tag{E.25}$$

$$= \mathrm{Im}(\theta_\alpha) \sum_{k=1}^{L} \lambda_k^{(\alpha)} \langle \phi_k | \sum_{j=1}^{L} d_{jk}^{(SA)} \hat{X}_{jk} \hat{H} \hat{X}_\alpha | \phi_k \rangle \tag{E.26}$$

$$= \mathrm{Im}(\theta_\alpha) \sum_{k=1}^{L} \lambda_k^{(\alpha)} \langle \phi_k | \hat{H}_k^{(SA)} \hat{X}_\alpha | \phi_k \rangle , \tag{E.27}$$

where in the second equality we have defined $d_{jk}^{(SA)} \equiv \sum_{I=1}^{N_s} w_I c_k^{(I)} c_j^{(I)}$, and in the third equality we have defined $\hat{H}_k^{(SA)} \equiv \sum_{j=1}^{L} d_{jk}^{(SA)} \hat{X}_{jk} \hat{H}$. Noting the resemblance between Eq.(E.27) and Eq.(E.10), we can straightforwardly apply the same techniques to obtain the state-averaged DIS $\mathcal{D}^{(SA)}$ as were used in Section E.1 to obtain $\mathcal{D}^{(I)}$, by replacing the $\hat{H}_k^{(I)}$'s with $\hat{H}_k^{(SA)}$'s.

Appendix F

Phase alignment procedures

For a multiconfigurational reference state, or a state-averaged mixture of states (whether they be single Slater determinants, or linear combinations thereof in the multiconfigurational case), the energy gradient magnitude expression for generator $\hat{P}_\alpha$ takes the general form of

$$g_\alpha = |\operatorname{Im}(\theta_\alpha)| \left| \sum_{k=0}^{L-1} \lambda_k^{(\alpha)} \Xi_k^{(\alpha)} \right|, \tag{F.1}$$

recalling that $\vec{\boldsymbol{\Xi}}^{(\alpha)} = \left\{ \Xi_0^{(\alpha)}, \Xi_1^{(\alpha)}, \ldots, \Xi_{L-1}^{(\alpha)} \right\}$ is fully determined by choice of $\hat{X}_\alpha$ ($\vec{\mu}^{(\alpha)}$). Once $\mathcal{D}^{(I/SA)}$ has been obtained, we wish to find the $\hat{P}_\alpha$ possessing the highest gradient g_α. In this Appendix, we describe how to select $\hat{Z}_\alpha$ which maximizes g_α, for a given $\hat{X}_\alpha$. That is, for fixed $\vec{\boldsymbol{\Xi}}^{(\alpha)}$, we aim to find $\hat{Z}_\alpha$ which produces maximal g_α through the $\hat{Z}_\alpha$ dependence of Eq.(F.1) in $\operatorname{Im}(\theta_\alpha)$ and $\{\lambda_k^{(\alpha)}\}_{k=0}^{L-1}$.

First, we formulate this problem in the domain of $\hat{Z}_\alpha$'s binary representation, $\vec{\nu}^{(\alpha)}$. We then describe the optimal strategy to finding $\vec{\nu}^{(\alpha)}$ which maximizes g_α in Appendix F.0.1, and an efficient yet heuristic strategies in Appendices F.0.2 and F.0.3.

Recalling that $g_\alpha = 0$ unless $\operatorname{Im}(\theta_\alpha) \in \{1, -1\}$, which is the case when $\hat{X}_\alpha$ and $\hat{Z}_\alpha$ have odd-valued overlapping support, leading to an odd number of $\hat{y}$ instances in $\hat{P}_\alpha$. In terms of the binary vectors, this leads to requirement

$$\vec{\mu}^{(\alpha)} \cdot \vec{\nu}^{(\alpha)} \mod 2 = 1. \tag{F.2}$$

The relative phases $\{\lambda_k^{(\alpha)}\}_{k=0}^{L-1}$ arise from expectation values $\{\langle\phi_k| \hat{Z}_\alpha |\phi_k\rangle\}_{k=0}^{L-1}$, where

$$\lambda_k^{(\alpha)} = \langle\phi_k| \hat{Z}_\alpha |\phi_k\rangle = \prod_{p=1}^{N} \left\langle \phi_k^{(p)} \right| \hat{z}_\alpha^{\nu_p^{(\alpha)}} \left| \phi_k^{(p)} \right\rangle \in \{1, -1\}, \tag{F.3}$$

where $\left|\phi_k^{(p)}\right\rangle \in \{|0\rangle, |1\rangle\}$, and $\hat{Z}_\alpha$ is represented using Eq.(2.16). Let us introduce an N_q-bit

binary vector $\vec{\phi}^{(k)} = (\phi_1^{(k)}, \dots \phi_N^{(k)})$ representing the computational basis state $|\phi_k\rangle$, that is

$$\vec{\phi}_p^{(k)} = \begin{cases} 1 \text{ if } \left|\phi_k^{(p)}\right\rangle = |1\rangle \\ 0 \text{ if } \left|\phi_k^{(p)}\right\rangle = |0\rangle \end{cases} . \tag{F.4}$$

Eigenvalue $\lambda_k^{(\alpha)} = -1$ if and only if there are an odd number of instances where $\mu_p^{(\alpha)} = \phi_p^{(k)} = 1$, otherwise $\lambda_k^{(\alpha)} = 1$. Hence, $\lambda_k^{(\alpha)}$ is written as a function of binary vectors $\vec{\mu}^{(\alpha)}$ and $\vec{\phi}^{(k)}$ as

$$\lambda_k^{(\alpha)} = 1 - 2\left(\vec{\phi}^{(k)} \cdot \vec{\mu}^{(\alpha)} \mod 2\right). \tag{F.5}$$

F.0.1 Optimal strategy

Given an element $\left(\vec{\mu}^{(\alpha)}, \vec{\vec{\Xi}}^{(\alpha)}\right) \in \mathcal{D}^{(I/SA)}$, maximizing g_α [Eq.(F.1)] through $\vec{\nu}^{(\alpha)}$ can be formulated as an N_q bit constrained binary optimization:

$$\begin{aligned} &\max_{\vec{\nu}^{(\alpha)}} C(\vec{\nu}^{(\alpha)}) \\ &\text{subject to: } \vec{\mu}^{(\alpha)} \cdot \vec{\nu}^{(\alpha)} \mod 2 = 1 \end{aligned} \tag{F.6}$$

where

$$C(\vec{\nu}^{(\alpha)}) = \left| \sum_{k=0}^{L-1} \Xi_k^{(\alpha)} \left[1 - 2\left(\vec{\phi}^{(k)} \cdot \vec{\mu}^{(\alpha)} \mod 2\right)\right] \right|. \tag{F.7}$$

Noting that $1 - 2\left(\vec{\phi}^{(k)} \cdot \vec{\mu}^{(\alpha)} \mod 2\right)$ is equivalent to Eq.(F.3), Eq.(F.7) consists of L distinct N_q-bit clauses, and hence it is generally exponentially hard to find the optimal value of $\vec{\nu}^{(\alpha)}$. Noting that $\vec{\vec{\Xi}}^{(\alpha)}$ has been precomputed, evaluation of Eq.(F.7) is computationally fast, with scaling $O(LN)$. When brute force searches are implausible for sufficiently large N_q, heuristic binary optimization strategies involving local searches may be feasible, such as e.g., Bayesian optimization.

F.0.2 Efficient greedy strategy

Since finding the optimal solution to Eq.(F.7) can be exponentially hard, we provide one heuristic algorithm for finding approximately optimal solutions herein. As it turns out, one can formulate whether the *idealized* scenario, if there exists a $\vec{\nu}^{(\alpha)}$ which aligns *all* signs of terms $\{\lambda_k^{(\alpha)}\Xi_k^{(\alpha)}\}_{k=0}^{L-1}$ in Eq.(F.7), as an efficiently solvable satisfiability problem. If satisfied, the algorithm returns the satisfying $\vec{\nu}^{(\alpha)}$. If unsatisfiable, one can remove constraints existing in the problem and check again for satisfiability. We refer to this procedure as the `GreedySAT` phase alignment routine, and explain it in further detail below.

If the sign of all terms $\{\lambda_k^{(\alpha)}\Xi_k^{(\alpha)}\}_{k=0}^{L-1}$ can be made uniform through $\lambda_k^{(\alpha)}$'s, and if $\mathrm{Im}(\theta_\alpha) \in \{1, -1\}$, then Eq.(F.1) can be simplified to

$$g_\alpha = \sum_{k=0}^{L-1} \left|\Xi_k^{(\alpha)}\right|. \tag{F.8}$$

Since a given $\Xi_k^{(\alpha)}$ can be of either sign, the goal is to find a $\hat{Z}_\alpha$ $(\vec{\nu}^{(\alpha)})$ which 'aligns' the signs of all $\lambda_k^{(\alpha)}\Xi_k^{(\alpha)}$ terms, recalling that $\hat{Z}_\alpha |\phi_k\rangle = \lambda_k^{(\alpha)} |\phi_k\rangle$. In other words, we require that for all nonzero $\Xi_k^{(\alpha)}$, $\Xi_l^{(\alpha)}$,

$$\mathrm{sgn}(\lambda_k^{(\alpha)}\Xi_k^{(\alpha)}) = \mathrm{sgn}(\lambda_l^{(\alpha)}\Xi_l^{(\alpha)}), \tag{F.9}$$

where $\mathrm{sgn}(x)$ is the signum function for $x \in \mathbb{R}$,

$$\mathrm{sgn}(x) = \begin{cases} 1 \text{ if } x > 0 \\ 0 \text{ if } x = 0 \\ -1 \text{ if } x < 0. \end{cases} \tag{F.10}$$

To satisfy Eq.(F.9), two satisfactory assignments exist,

$$\lambda_k^{(\alpha)} = (\pm)\mathrm{sgn}\left(\Xi_k^{(\alpha)}\right), \tag{F.11}$$

for all considered $\Xi_k^{(\alpha)}$, where choice of $(\pm)$ is held fixed for all k . The freedom of choosing $(\pm)$ comes from the fact that to satisfy Eq.(F.8), terms in Eq.(F.1) can be either be all positive or all negative.

Attempting to find $\vec{\nu}^{(\alpha)}$ which satisfies Eq.(F.11) for all $L' \leq L$ non-zero $\Xi_k^{(\alpha)}$'s, along with the requirement that $\vec{\mu}^{(\alpha)} \cdot \vec{\nu}^{(\alpha)} \mod 2 = 1$, leads to a system of $L' + 1$ equations on the N_q binary variables in $\vec{\nu}^{(\alpha)}$. We can reformulate the conditions of Eq.(F.11) as

$$\vec{\phi}^{(k)} \cdot \vec{\nu}^{(\alpha)} = \frac{(\pm)\mathrm{sgn}(\Xi_k^{(\alpha)}) - 1}{2} \mod 2. \tag{F.12}$$

We can then formulate the system of equations as

$$\boldsymbol{M}\vec{\nu}^{(\alpha)} = \vec{\boldsymbol{b}}, \tag{F.13}$$

where $\boldsymbol{M}$ is a $(L'+1) \times N_q$ matrix,

$$\boldsymbol{M} = \begin{pmatrix} \vec{\mu}_1^{(\alpha)} & \dots & \vec{\mu}_{N_q}^{(\alpha)} \\ \vec{\phi}_1^{(1)} & \dots & \vec{\phi}_{N_q}^{(1)} \\ & \vdots & \\ \vec{\phi}_1^{(L')} & \dots & \vec{\phi}_{N_q}^{(L')} \end{pmatrix}, \tag{F.14}$$

and $\vec{\boldsymbol{b}}$ is a $L'+1$ dimensional vector,

$$\vec{\boldsymbol{b}} = \begin{pmatrix} 1 \\ \left[(\pm)\text{sgn}(\Xi_1^{(\alpha)}) - 1\right]/2 \mod 2 \\ \vdots \\ \left[(\pm)\text{sgn}(\Xi_{L'}^{(\alpha)}) - 1\right]/2 \mod 2 \end{pmatrix}. \tag{F.15}$$

We can then solve Eq.(F.13) by binary Gaussian elimination, for instance.

We now describe the complete `GreedySAT` procedure below, which includes a prescription for when no solution to Eq.(F.13) can be found. Essentially, if no solution can be found for considering the phase alignment of all L' terms via satisfying all Eq.(F.12), we remove consideration of the specific instance of Eq.(F.12) associated with the lowest valued $|\Xi_k^{(\alpha)}|$, and attempt to solve the system of fewer equations. This removal of least-important constraints is iteratively performed until a solution $\vec{\nu}^{(\alpha)}$ has been found, and the true value of the associated g_α is computed via Eq.(F.7).

1. For $\left(\vec{\mu}^{(\alpha)}, \vec{\boldsymbol{\Xi}}^{(\alpha)}\right) \in \mathcal{D}^{(I/SA)}$, let $\vec{\boldsymbol{\Xi}}_{\text{nz}}^{(\alpha)}$ denote the $L' \leq L$ dimensional vector of the non-zero components of $\vec{\boldsymbol{\Xi}}^{(\alpha)}$, and attempt to find solution to Eq.(F.13) for the $(+)$ assignment in Eq.(F.11) for all $\Xi_k^{(\alpha)} \in \vec{\boldsymbol{\Xi}}_{\text{nz}}^{(\alpha)}$. If no solution exists, attempt to solve the same system but with the $(-)$ assignment in Eq.(F.11). If a solution has been found for either $(\pm)$ assignments, go to Step 3 with solution $\vec{\nu}^{(\alpha)}$ to Eq.(F.13), otherwise, enter Step 2.

2. If no solution was found in Step 1, we prune the consideration of the $\vec{\phi}^{(k)}$ with lowest $|\Xi_k^{(\alpha)}|$, i.e., obtain $\tilde{\boldsymbol{\Xi}}_{\text{nz}}^{(\alpha)}$ of dimension $L'' = L' - 1$, where the lowest component by absolute value has been removed. Enter Step 1 using $\tilde{\boldsymbol{\Xi}}_{\text{nz}}^{(\alpha)}$.

3. Once a solution $\vec{\nu}^{(\alpha)}$ has been found, the corresponding Pauli term $\hat{P}_\alpha$ has $\hat{X}_\alpha$ and $\hat{Z}_\alpha$ parts given by $\vec{\mu}^{(\alpha)}$ and $\vec{\nu}^{(\alpha)}$ respectively, and it's gradient magnitude g_α is given by inserting the found $\vec{\nu}^{(\alpha)}$ into Eq.(F.7). This resulting $\hat{P}_\alpha$ represents the highest g_α candidate with $\hat{X}_\alpha$ specified by $\vec{\mu}^{(\alpha)}$ found by the `GreedySAT` routine.

F.0.3 Continuous relaxation

Herein, we discuss another route to approximately finding the optimal $\vec{\nu}^{(\alpha)}$ for Eq.(F.7) involving a relaxation from N_q binary variables, to a set of N_q continuous variables in the interval $[0,1]$. To circumvent binary optimization, one can replace $C(\vec{\nu}^{(\alpha)})$ with a cost function of continuous variables, $\tilde{C}(\vec{v}^{(\alpha)})$, where $\vec{v}^{(\alpha)} = (v_1^{(\alpha)}, \ldots, v_{N_q}^{(\alpha)})$ with $v_p \in [0,1]$,

$$\tilde{C}(\vec{v}^{(\alpha)}) = \left| \sum_{k=0}^{L-1} \Xi_k^{(\alpha)} f_k(\vec{v}^{(\alpha)}) \right|, \tag{F.16}$$

where $\lambda_k^{(\alpha)}$'s have been replaced by continuous functions $f_k : \mathbb{R}^{N_q} \to \mathbb{R}$, with image in the interval $[-1,1]$. Essentially, rather than having a discrete set of variables $\vec{\nu}^{(\alpha)}$ dictate the binary placements of $\hat{z}_p$'s in $\hat{Z}_\alpha$, we can, for purpose of optimization, utilize continuous set of variables $\vec{v}^{(\alpha)}$, which dictate the degree of single-qubit depolarizations applied to a N_q-qubit product of $\hat{z}_p$ operators,

$$f_k(\vec{v}^{(\alpha)}) = \langle \phi_k | \prod_{p=1}^{N_q} \Delta_{v_p^{(\alpha)}}(\hat{z}_p) \, | \phi_k \rangle \,, \tag{F.17}$$

where

$$\Delta_{v_p^{(\alpha)}}(\hat{z}_p) = v_p^{(\alpha)} \hat{z}_p + (1 - v_p^{(\alpha)}) \hat{I} \tag{F.18}$$

and Eq.(F.17) can be explicitly written as

$$f_k(\vec{v}^{(\alpha)}) = \prod_{p=1}^{N_q} \left(v_p^{(\alpha)} \left\langle \phi_k^{(p)} \right| \hat{z}_p \left| \phi_k^{(p)} \right\rangle + (1 - v_p^{(\alpha)}) \right). \tag{F.19}$$

Under such a parameterization, Eq.(F.16) can then be optimized via, e.g., gradient-descent algorithms. However, to arrive at a final $\hat{Z}_\alpha$ which corresponds to an approximately maximizing $\vec{\nu}^{(\alpha)}$ of C, one must employ a *rounding procedure* which takes optimal $\vec{v}^{(\alpha)*} = \text{argmax}_{\vec{v}^{(\alpha)}} \tilde{C}$ to a binary vector, $\vec{v}^{(\alpha)*} \to \vec{\nu}^{(\alpha)*}$. Furthermore, the rounding procedure must ensure that $\vec{\nu}^{(\alpha)*}$ satisfies $\vec{\mu}^\alpha \cdot \vec{\nu}^{(\alpha)*} \mod 2 = 1$.

Bibliography

[1]I. G. Ryabinkin, R. A. Lang, S. N. Genin, and A. F. Izmaylov, "Iterative qubit coupled cluster approach with efficient screening of generators", J. Chem. Theory Comput. **16**, 1055–1063 (2020).

[2]R. A. Lang, I. G. Ryabinkin, and A. F. Izmaylov, "Unitary transformation of the electronic Hamiltonian with an exact quadratic truncation of the Baker-Campbell-Hausdorff expansion", J. Chem. Theory Comput. **17**, 66–78 (2021).

[3]R. A. Lang, A. Ganeshram, and A. F. Izmaylov, "Growth reduction of similarity-transformed electronic Hamiltonians in qubit space", J. Chem. Theory Comput. **19**, 6656–6667 (2023).

[4]R. A. Lang, I. G. Ryabinkin, and A. F. Izmaylov, "A multireference state-averaged iterative qubit coupled cluster method for ground and excited state electronic energies", Under preparation.

[5]S. N. Genin, I. G. Ryabinkin, N. R. Paisley, S. O. Whelan, M. G. Helander, and Z. M. Hudson, "Estimating phosphorescent emission energies in Ir^{III} complexes using large-scale quantum computing simulations", Angew. Chem. Int. Ed. **61**, e202116175 (2022).

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

[7]G. Hao Low and N. Wiebe, "Hamiltonian Simulation in the Interaction Picture", arXiv:1805.00675 (2018).

[8]Y. Su, D. W. Berry, N. Wiebe, N. Rubin, and R. Babbush, "Fault-tolerant quantum simulations of chemistry in first quantization", PRX Quantum **2**, 040332 (2021).

[9]Q. Sun, X. Zhang, S. Banerjee, P. Bao, M. Barbry, N. S. Blunt, N. A. Bogdanov, G. H. Booth, J. Chen, Z.-H. Cui, J. J. Eriksen, Y. Gao, S. Guo, J. Hermann, M. R. Hermes, K. Koh, P. Koval, S. Lehtola, Z. Li, J. Liu, N. Mardirossian, J. D. McClain, M. Motta, B. Mussard, H. Q. Pham, A. Pulkin, W. Purwanto, P. J. Robinson, E. Ronca, E. R. Sayfutyarova, M. Scheurer, H. F. Schurkus, J. E. T. Smith, C. Sun, S.-N. Sun, S. Upadhyay, L. K. Wagner, X. Wang, A. White, J. D. Whitfield, M. J. Williamson, S. Wouters, J. Yang, J. M. Yu, T. Zhu, T. C. Berkelbach, S. Sharma, A. Y. Sokolov, and G. K.-L. Chan, "Recent developments in the PySCF program package", J. Chem. Phys. **153**, 024109 (2020).

[10]W. Kutzelnigg, "Separation of strong (bond-breaking) from weak (dynamical) correlation", Chem. Phys. **401**, 119–124 (2012).

[11]D. Hait, N. M. Tubman, D. S. Levine, K. B. Whaley, and M. Head-Gordon, "What levels of coupled cluster theory are appropriate for transition metal systems? a study using near-exact quantum chemical values for 3d transition metal binary compounds", J. Chem. Theory Comput. **15**, 5370–5385 (2019).

[12]J. Preskill, "Quantum Computing in the NISQ era and beyond", Quantum **2**, 79 (2018).

[13]Y. Cao, J. Romero, J. P. Olson, M. Degroote, P. D. Johnson, M. Kieferová, I. D. Kivlichan, T. Menke, B. Peropadre, N. P. D. Sawaya, S. Sim, L. Veis, and A. Aspuru-Guzik, "Quantum chemistry in the age of quantum computing", Chem. Rev. **119**, 10856–10915 (2019).

[14]J. R. McClean, J. Romero, R. Babbush, and A. Aspuru-Guzik, "The theory of variational hybrid quantum-classical algorithms", New J. Phys. **18**, 023023 (2016).

[15]V. Verteletskyi, T.-C. Yen, and A. F. Izmaylov, "Measurement optimization in the variational quantum eigensolver using a minimum clique cover", J.Chem. Phys. **152**, 124114 (2020).

[16]A. F. Izmaylov, T.-C. Yen, R. A. Lang, and V. Verteletskyi, "Unitary partitioning approach to the measurement problem in the variational quantum eigensolver method", J. Chem. Theory Comput. **16**, 190–195 (2020).

[17]T.-C. Yen and A. F. Izmaylov, "Cartan subalgebra approach to efficient measurements of quantum observables", PRX Quantum **2**, 040320 (2021).

[18]P. Jordan and E. P. Wigner, "Über das paulische äquivalenzverbot", Eur. Phys. J. A **47**, 631–651 (1928).

[19]A. Aspuru-Guzik, A. D. Dutoi, P. J. Love, and M. Head-Gordon, "Simulated quantum computation of molecular energies", Science **309**, 1704–1707 (2005).

[20]S. B. Bravyi and A. Y. Kitaev, "Fermionic quantum computation", Ann. Phys. **298**, 210–226 (2002).

[21]J. T. Seeley, M. J. Richard, and P. J. Love, "The Bravyi-Kitaev transformation for quantum computation of electronic structure", J. Chem. Phys. **137**, 224109 (2012).

[22]A. Tranter, P. J. Love, F. Mintert, and P. V. Coveney, "A comparison of the bravyi–kitaev and jordan–wigner transformations for the quantum simulation of quantum chemistry", J. Chem. Theory Comput. **14**, 5617–5630 (2018).

[23]A. Aspuru-Guzik, A. D. Dutoi, P. J. Love, and M. Head-Gordon, "Simulated quantum computation of molecular energies", Science **309**, 1704–1707 (2005).

[24]R. B. Griffiths and C.-S. Niu, "Semiclassical fourier transform for quantum computation", Phys. Rev. Lett. **76**, 3228–3231 (1996).

[25]L. Lin and Y. Tong, "Heisenberg-limited ground-state energy estimation for early fault-tolerant quantum computers", PRX Quantum **3**, 010318 (2022).

[26]G. Wang, D. S. França, R. Zhang, S. Zhu, and P. D. Johnson, *Quantum algorithm for ground state energy estimation using circuit depth with exponentially improved dependence on precision*, 2023.

[27]Z. Ding and L. Lin, "Even shorter quantum circuit for phase estimation on early fault-tolerant quantum computers with applications to ground-state energy estimation", PRX Quantum **4**, 020331 (2023).

[28]T.-C. Yen, V. Verteletskyi, and A. F. Izmaylov, "Measuring all compatible operators in one series of single-qubit measurements using unitary transformations", J. Chem. Theory Comput. **16**, 2400–2409 (2020).

[29]R. H. Byrd, P. Lu, J. Nocedal, and C. Zhu, "A limited memory algorithm for bound constrained optimization", SIAM J. Sci. Comput. **16**, 1190–1208 (1995).

[30]A. R. Conn, K. Scheinberg, and P. L. Toint, "On the convergence of derivative-free methods for unconstrained optimization", Approximation theory and optimization: tributes to MJD Powell, 83–108 (1997).

[31]D. M. Olsson and L. S. Nelson, "The Nelder-Mead simplex procedure for function minimization", Technometrics **17**, 45–51 (1975).

[32]J. Stokes, J. Izaac, N. Killoran, and G. Carleo, "Quantum Natural Gradient", Quantum **4**, 269 (2020).

[33]J. Gacon, C. Zoufal, G. Carleo, and S. Woerner, "Simultaneous Perturbation Stochastic Approximation of the Quantum Fisher Information", Quantum **5**, 567 (2021).

[34]A. Peruzzo, J. McClean, P. Shadbolt, M.-H. Yung, X.-Q. Zhou, P. J. Love, A. Aspuru-Guzik, and J. L. O'Brien, "A variational eigenvalue solver on a photonic quantum processor", Nat. Commun. **5**, 4213 (2014).

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

[36]J. Romero, R. Babbush, J. R. McClean, C. Hempel, P. J. Love, and A. Aspuru-Guzik, "Strategies for quantum computing molecular energies using the unitary coupled cluster ansatz", Quantum Sci. Technol. **4**, 014008 (2018).

[37]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, "Quantum Chemistry Calculations on a Trapped-Ion Quantum Simulator", Phys. Rev. X **8**, 31022 (2018).

[38]Y. Nam, J.-S. Chen, N. C. Pisenti, K. Wright, C. Delaney, D. Maslov, K. R. Brown, S. Allen, J. M. Amini, J. Apisdorf, K. M. Beck, A. Blinov, V. Chaplin, M. Chmielewski, C. Collins, S. Debnath, K. M. Hudek, A. M. Ducore, M. Keesan, S. M. Kreikemeier, J. Mizrahi, P. Solomon, M. Williams, J. D. Wong-Campos, D. Moehring, C. Monroe, and J. Kim, "Ground-state energy estimation of the water molecule on a trapped-ion quantum computer", npj Quantum Inf. **6**, 33 (2020).

[39]J. Chen, H.-P. Cheng, and J. K. Freericks, "Quantum-inspired algorithm for the factorized form of unitary coupled cluster theory", J. Chem. Theory Comput. **17**, 841–847 (2021).

[40]A. F. Izmaylov, M. Díaz-Tinoco, and R. A. Lang, "On the order problem in construction of unitary operators for the variational quantum eigensolver", Phys. Chem. Chem. Phys. **22**, 12980–12986 (2020).

[41]F. A. Evangelista, G. K.-L. Chan, and G. E. Scuseria, "Exact parameterization of fermionic wave functions via unitary coupled cluster theory", J. Chem. Phys. **151**, 244112 (2019).

[42]H. R. Grimsley, D. Claudino, S. E. Economou, E. Barnes, and N. J. Mayhall, "Is the trotterized UCCSD ansatz chemically well-defined?", J. Chem. Theory Comput. **16**, 1–6 (2020).

[43]J. R. McClean, S. Boixo, V. N. Smelyanskiy, R. Babbush, and H. Neven, "Barren plateaus in quantum neural network training landscapes", Nat. Commun. **9**, 1–6 (2018).

[44]M. Cerezo, A. Sone, T. Volkoff, L. Cincio, and P. J. Coles, "Cost function dependent barren plateaus in shallow parametrized quantum circuits", Nat. Commun. **12**, 1791 (2021).

[45]H. R. Grimsley, S. E. Economou, E. Barnes, and N. J. Mayhall, "An adaptive variational algorithm for exact molecular simulations on a quantum computer", Nat. Commun. **10**, 3007 (2019).

[46]I. G. Ryabinkin, T.-C. Yen, S. N. Genin, and A. F. Izmaylov, "Qubit coupled cluster method: a systematic approach to quantum chemistry on a quantum computer", J. Chem. Theory Comput. **14**, 6317–6326 (2018).

[47]Y. S. Yordanov, V. Armaos, C. H. W. Barnes, and D. R. M. Arvidsson-Shukur, "Qubit-excitation-based adaptive variational quantum eigensolver", Commun. Phys. **4**, 228 (2021).

[48]H. L. Tang, V. Shkolnikov, G. S. Barron, H. R. Grimsley, N. J. Mayhall, E. Barnes, and S. E. Economou, "Qubit-adapt-vqe: an adaptive algorithm for constructing hardware-efficient ansätze on a quantum processor", PRX Quantum **2**, 020310 (2021).

[49]J. F. Gonthier, M. D. Radin, C. Buda, E. J. Doskocil, C. M. Abuan, and J. Romero, "Measurements as a roadblock to near-term practical quantum advantage in chemistry: resource analysis", Phys. Rev. Res. **4**, 033154 (2022).

[50]J. R. McClean, R. Babbush, P. J. Love, and A. Aspuru-Guzik, "Exploiting locality in quantum computation for quantum chemistry", J. Phys. Chem. Lett. **5**, 4368–4380 (2014).

[51]T.-C. Yen, A. Ganeshram, and A. F. Izmaylov, "Deterministic improvements of quantum measurements with grouping of compatible operators, non-local transformations, and covariance estimates", npj Quantum Inf **9**, 14 (2023).

[52]S. Choi, I. Loaiza, and A. F. Izmaylov, "Fluid fermionic fragments for optimizing quantum measurements of electronic Hamiltonians in the variational quantum eigensolver", Quantum **7**, 889 (2023).

[53]J. A. Pople, "Nobel lecture: quantum chemical models", Rev. Mod. Phys. **71**, 1267–1274 (1999).

[54]K. A. Peterson, D. Feller, and D. A. Dixon, "Chemical accuracy in ab initio thermochemistry and spectroscopy: current strategies and future challenges", Theor. Chem. Acc. **131**, 1079 (2012).

[55]J. S. Kottmann, A. Anand, and A. Aspuru-Guzik, "A feasible approach for automatically differentiable unitary coupled-cluster on quantum computers", Chem. Sci. **12**, 3497–3508 (2021).

[56]I. G. Ryabinkin, S. N. Genin, and A. F. Izmaylov, "Constrained variational quantum eigensolver: quantum computer search engine in the fock space", J. Chem. Theory Comput. **15**, 249–255 (2019).

[57]I. G. Ryabinkin, S. N. Genin, and A. F. Izmaylov, "Relation between fermionic and qubit mean fields in the electronic structure problem", J. Chem. Phys. **149**, 214105 (2018).

[58]J. Li, X. Yang, X. Peng, and C.-P. Sun, "Hybrid quantum-classical approach to quantum optimal control", Phys. Rev. Lett. **118**, 150503 (2017).

[59]M. Schuld, V. Bergholm, C. Gogolin, J. Izaac, and N. Killoran, "Evaluating analytic gradients on quantum hardware", Phys. Rev. A **99**, 032331 (2019).

[60]A. F. Izmaylov, R. A. Lang, and T.-C. Yen, "Analytic gradients in variational quantum algorithms: algebraic extensions of the parameter-shift rule to general unitary transformations", Phys. Rev. A **104**, 062443 (2021).

[61]J. Liu, Z. Li, and J. Yang, "Reducing circuit depth in adaptive variational quantum algorithms via effective hamiltonian theories", J. Chem. Theory Comput. **18**, 4795–4805 (2022).

[62]N. P. Bauman, E. J. Bylaska, S. Krishnamoorthy, G. H. Low, N. Wiebe, C. E. Granade, M. Roetteler, M. Troyer, and K. Kowalski, "Downfolding of many-body hamiltonians using active-space models: extension of the sub-system embedding sub-algebras approach to unitary coupled cluster formalisms", J. Chem. Phys. **151**, 014107 (2019).

[63]M. Metcalf, N. P. Bauman, K. Kowalski, and W. A. de Jong, "Resource-efficient chemistry on quantum computers with the variational quantum eigensolver and the double unitary coupled-cluster approach", J. Chem. Theory Comput. **16**, 6165–6175 (2020).

[64]M. Motta, T. P. Gujarati, J. E. Rice, A. Kumar, C. Masteran, J. A. Latone, E. Lee, E. F. Valeev, and T. Y. Takeshita, "Quantum simulation of electronic structure with a transcorrelated Hamiltonian: improved accuracy with a smaller footprint on the quantum computer", Phys. Chem. Chem. Phys. **22**, 24270–24281 (2020).

[65]R. Huang, C. Li, and F. A. Evangelista, "Leveraging small-scale quantum computers with unitarily downfolded Hamiltonians", PRX Quantum **4**, 020313 (2023).

[66]R. Sarkar and E. van den Berg, "On sets of commuting and anticommuting Paulis", arXiv:1909.08123 (2019).

[67]S. Bravyi, J. M. Gambetta, A. Mezzacapo, and K. Temme, "Tapering off qubits to simulate fermionic Hamiltonians", arXiv:1701.08213 (2017).

[68]I. G. Ryabinkin, A. F. Izmaylov, and S. N. Genin, "A posteriori corrections to the iterative qubit coupled cluster method to minimize the use of quantum resources in large-scale calculations", Quantum Sci. Technol. **6**, 024012 (2021).

[69]J. M. Turney, A. C. Simmonett, R. M. Parrish, E. G. Hohenstein, F. A. Evangelista, J. T. Fermann, B. J. Mintz, L. A. Burns, J. J. Wilke, M. L. Abrams, N. J. Russ, M. L. Leininger, C. L. Janssen, E. T. Seidl, W. D. Allen, H. F. Schaefer, R. A. King, E. F. Valeev, C. D. Sherrill, and T. D. Crawford, "Psi4: an open-source ab initio electronic structure program", WIREs Comput. Mol. Sci. **2**, 556–565 (2012).

[70]J. S. Kottmann, S. Alperin-Lea, T. Tamayo-Mendoza, A. Cervera-Lierta, C. Lavigne, T.-C. Yen, V. Verteletskyi, P. Schleich, A. Anand, M. Degroote, S. Chaney, M. Kesibi, N. G. Curnow, B. Solo, G. Tsilimigkounakis, C. Zendejas-Morales, A. F. Izmaylov, and A. Aspuru-Guzik, "Tequila: a platform for rapid development of quantum algorithms", Quantum Sci. Technol. **6**, 024009 (2021).

[71]Qiskit contributors, *Qiskit: an open-source framework for quantum computing*, 2023.

[72]G. M. J. Barca, C. Bertoni, L. Carrington, D. Datta, N. De Silva, J. E. Deustua, D. G. Fedorov, J. R. Gour, A. O. Gunina, E. Guidez, T. Harville, S. Irle, J. Ivanic, K. Kowalski, S. S. Leang, H. Li, W. Li, J. J. Lutz, I. Magoulas, J. Mato, V. Mironov, H. Nakata, B. Q. Pham, P. Piecuch, D. Poole, S. R. Pruitt, A. P. Rendell, L. B. Roskop, K. Ruedenberg, T. Sattasathuchana, M. W. Schmidt, J. Shen, L. Slipchenko, M. Sosonkina, V. Sundriyal, A. Tiwari, J. L. Galvez Vallejo, B. Westheimer, M. Wloch, P. Xu, F. Zahariev, and M. S. Gordon, "Recent developments in the general atomic and molecular electronic structure system", J. Chem. Phys. **152**, 154102 (2020).

[73]V. E. Elfving, B. W. Broer, M. Webber, J. Gavartin, M. D. Halls, K. P. Lorton, and A. Bochevarov, "How will quantum computers provide an industrially relevant computational advantage in quantum chemistry?", arXiv:2009.12472 (2020).

[74]B. Bauer, S. Bravyi, M. Motta, and G. K.-L. Chan, "Quantum algorithms for quantum chemistry and quantum materials science", Chem. Rev. **120**, 12685–12717 (2020).

[75]S. Lee, J. Lee, H. Zhai, Y. Tong, A. M. Dalzell, A. Kumar, P. Helms, J. Gray, Z.-H. Cui, W. Liu, M. Kastoryano, R. Babbush, J. Preskill, D. R. Reichman, E. T. Campbell, E. F. Valeev, L. Lin, and G. K.-L. Chan, "Evaluating the evidence for exponential quantum advantage in ground-state quantum chemistry", Nat. Commun. **14**, 1952 (2023).

[76]A. Zhao, A. Tranter, W. M. Kirby, S. F. Ung, A. Miyake, and P. J. Love, "Measurement reduction in variational quantum algorithms", Phys. Rev. A **101**, 062322 (2020).

[77]M. Mézard and A. Montanari, *Information, physics, and computation*, Oxford Graduate Texts (OUP Oxford, 2009).

[78]S. N. Genin, I. G. Ryabinkin, and A. F. Izmaylov, "Quantum chemistry on quantum annealers", arXiv:1901.04715 (2019).