Advancing Near-Term Quantum Computing for Chemistry: Multireference-State Error Mitigation

Master's thesis in Nanotechnology

Hang Zou

DEPARTMENT OF CHEMISTRY AND CHEMICAL ENGINEERING

CHALMERS UNIVERSITY OF TECHNOLOGY
Gothenburg, Sweden 2024
www.chalmers.se

Master's thesis 2024

Advancing Near-Term Quantum Computing for Chemistry: Multireference-State Error Mitigation

Hang Zou

CHALMERS
UNIVERSITY OF TECHNOLOGY

Department of Chemistry and Chemical Engineering
Division of Physical Chemistry
Rahmlab group
Chalmers University of Technology
Gothenburg, Sweden 2024

Advancing Near-Term Quantum Computing for Chemistry: Multireference-State Error Mitigation
Hang Zou

Supervisor:Werner Dobrautz, Department of Chemistry and Chemical Engineering
Examiner: Martin Rahm, Department of Chemistry and Chemical Engineering

Master's Thesis 2024
Department of Chemistry and Chemical Engineering
Division of Physical Chemistry
Rahmlab group
Chalmers University of Technology
SE-412 96 Gothenburg
Telephone +46 31 772 1000

Typeset in LaTeX
Printed by Chalmers Reproservice
Gothenburg, Sweden 2024

Advancing Near-Term Quantum Computing for Chemistry: Multireference-State Error Mitigation

Hang Zou
Department of Chemistry and Chemical Engineering
Chalmers University of Technology

Abstract

Quantum error mitigation strategies are essential for enhancing the accuracy and reliability of quantum chemistry algorithms on noisy intermediate-scale quantum devices. Among these methods, Reference-state error mitigation (REM) stands out for its cost-effectiveness and chemical inspiration. However, the efficacy of REM is often limited when dealing with strongly correlated molecules. In this work, we introduce multireference-state error mitigation (MR-REM), an approach that builds upon the foundation of REM but extends its ability to capture the effects of noise on the ground state by leveraging multireference states. Central to MR-REM is the utilization of Givens rotations, which facilitate the construction of circuits capable of generating multireference states while respecting the underlying symmetries of the system. These states, tailored to exhibit significant overlap with the target ground state, enable effective error mitigation in variational quantum eigensolver experiments. We demonstrate the effectiveness of MR-REM through comprehensive simulations on molecules (H_2, H_2O, N_2, and F_2), highlighting its ability to achieve substantial improvements in computational accuracy compared to the original REM method. MR-REM extends the scope of error mitigation to encompass a wider range of molecular systems, including those with significant electronic correlation.

Keywords: Electronic structure, Quantum computation, VQE, Quantum error mitigation, Multireference.

Acknowledgements

I sincerely thank my examiner Martin Rahm and supervisor Werner Dobrautz for their selfless support and professional guidance throughout this research. Their profound knowledge and unwavering dedication have been instrumental in shaping the outcomes of this study. Furthermore, I extend our gratitude to Erika and Hampus, along with other team members, for their valuable discussions, which have greatly facilitated the smooth progress of our research work. Lastly, I wish to express our heartfelt appreciation to all my family and friends for their support and encouragement. Your belief in me has been a constant source of motivation throughout this journey.

Hang Zou, Gothenburg, June 2024

List of Acronyms

Below is the list of acronyms that have been used throughout this thesis listed in alphabetical order:

ADAPT-VQE	Adaptive Derivative-Assembled Pseudo-Trotter Ansatz Variational Quantum Eigensolver
BPs	Barren Plateaus
BK	Bravyi-Kitaev
CAS	Complete Active Space
CASSCF	Complete Active Space Self-Consistent Field
CCSD	Coupled-Cluster Singles and Doubles
CISD	Configuration Interaction Singles and Doubles
CPTP	Completely Positive Trace-Preserving
DMRG	Density Matrix Renormalization Group
FCI	Full Configuration Interaction
GTOs	Gaussian-Type Orbitals
HEA	Hardware-Efficient Ansatz
HF	Hartree-Fock
JW	Jordan-Wigner
k-UpCCGSD	k-Unitary Pair Coupled-Cluster Generalized Singles and Doubles
MCSCF	Multiconfiguration Self-Consistent Field
MR-REM	Multireference-State Error Mitigation
NISQ	Noisy Intermediate-Scale Quantum
PEC	Probabilistic Error Cancellation
PES	Potential Energy Surface
PQC	Parameterized Quantum Circuit
QAOA	Quantum Approximate Optimization Algorithm
QDT	Quantum Detector Tomography
QEC	Quantum Error Correction
QEM	Quantum Error Mitigation
QITE	Quantum Imaginary Time Evolution
QIT	Quantum Information Theory
QNP	Quantum Number Preserving
PES	Potential Energy Surface
QPE	Quantum Phase Estimation
REM	Reference-State Error Mitigation

SDs	Slater Determinants
tUPS	tiled Unitary Product States
UCC	Unitary Coupled Cluster
VQE	Variational Quantum Eigensolver
VQAs	Variational Quantum Algorithms
ZNE	Zero Noise Extrapolation

Contents

List of Acronyms viii

List of Figures xiii

List of Tables xv

1 Introduction 1
- 1.1 The age of quantum information . . . 1
- 1.2 Quantum chemistry on quantum computers . . . 4
- 1.3 Thesis outline . . . 5

2 Theory 7
- 2.1 The electronic structure problem . . . 7
  - 2.1.1 The molecular Hamiltonian . . . 7
  - 2.1.2 The Fermionic wavefunction . . . 8
    - 2.1.2.1 First quantization . . . 9
    - 2.1.2.2 Second quantization . . . 9
  - 2.1.3 Electron correlation . . . 10
  - 2.1.4 Active space . . . 11
- 2.2 Variational quantum eigensolver . . . 12
  - 2.2.1 Overview of VQE . . . 12
  - 2.2.2 Mapping to the qubit space . . . 13
    - 2.2.2.1 The Jordan-Wigner mapping . . . 13
    - 2.2.2.2 The parity mapping . . . 14
    - 2.2.2.3 The Bravyi-Kitaev mapping . . . 15
    - 2.2.2.4 Qubit tapering . . . 15
  - 2.2.3 Parameterized ansätze . . . 16
    - 2.2.3.1 Chemistry-inspired ansätze . . . 16
    - 2.2.3.2 Hardware-efficient ansätze . . . 17
    - 2.2.3.3 Symmetry-preserving ansätze . . . 18
- 2.3 Quantum error mitigation . . . 19
  - 2.3.1 Measurement error mitigation . . . 19
  - 2.3.2 Zero noise extrapolation . . . 19
  - 2.3.3 Probabilistic error cancellation . . . 20
  - 2.3.4 Symmetry verification by post-selection . . . 21
  - 2.3.5 Reference-state error mitigation . . . 21
- 2.4 Multireference-state error mitigation . . . 22

2.4.1 Motivation 22
2.4.2 The MR-REM procedure 23
2.4.3 Quantum lego: multireference state preparation 24

3 Methods **27**
3.1 Computational details 27
3.1.1 Noisy simulation 27
3.2 Resource statistics of quantum circuits 28

4 Results **31**
4.1 Potential energy surface 31
4.1.1 $H_2(2e, 2o)$ and $F_2(10e, 6o)$ 31
4.1.2 $H_2O(4e, 4o)$ and $N_2(6e, 6o)$ 32
4.2 Symmetry check 34

5 Conclusion **37**

Bibliography **39**

A Appendix **I**
A.1 Decompositions of Givens gates I
A.2 Convergence evaluation I

List of Figures

2.1 Determinants in a CAS(4e,4o) complete active space, where there are 4 active electrons in 4 active orbitals, and examples of single and double excitations. 12

2.2 Graphical representation of the workflow of VQE. 13

2.3 A strategy representation of the REM approach. Noise tends to flatten and elevate the energy landscape. Using a reference energy can partially capture the impact of noise and significantly mitigate the results of noisy VQE. 22

2.4 Multireference states approximate the ground state by considering a linear combination of multiple SDs. 23

2.5 Overview of MR-REM: comprising multireference state preparation and reference state error migration. Extending the preparation of HF state on a quantum computer to prepare multireference states using Givens gate-based circuits. 24

2.6 An example of $G(\theta)$ and $G^{(2)}(\theta)$ applied on a CAS(4e,4o) space. The qubits follow an interleaved spin ordering of electrons, i.e., $|\beta_3\alpha_3\beta_2\alpha_2\beta_1\alpha_1\beta_0\alpha_0\rangle$, where the rightmost qubit refers to the first qubit 0. $G(\theta)$ acts on qubits 4 and 2, achieving a single excitation of α_1 to α_2. $G^{(2)}(\theta)$ acts on qubits $4, 5$ and $2, 3$, achieving a double excitation of α_1, β_1 to α_2, β_2. 25

3.1 Schematic diagram of the complete circuit structure of MR-REM. All optimizable parameters $\boldsymbol{\theta}$ in the hardware-efficient R_Y ansatz are initialized to 0. The parameters $\boldsymbol{\phi}$ in the MR preparation circuit are obtained from a trial wavefunction through classical theory, which can be fixed or adjustable. 28

4.1 Comparisons of REM-2SDs and REM-HF for the potential energy surfaces (PES) of $H_2(2e, 2o)$ and $F_2(10e, 6o)$ at the cc-pVDZ level. The multireference 2SDs is obtained from truncated wavefunctions using the CISD method using a minimal STO-6G basis set. Left: PES for exact results, noisy VQE results with reference states of HF and 2SDs, and the corresponding REM results individually. Right: The absolute error of VQE and REM results with respect to the exact results. The cyan dashed lines represent the computational accuracy below 1.6×10^{-3} hartree (1 kcal/mol). 32

4.2 Comparison of REM-3SDs and REM-HF for the PES of $H_2O(4e, 4o)$ at the cc-pVDZ level. The PES scan varies both HO bond lengths R simultaneously, with a fixed $\angle HOH = 104.5°$. The multireference state of 3SDs is guided by CCSD/6-31G. 33
4.3 Comparison of REM-3SDs and REM-HF for the PES of $N_2(6e, 6o)$ at the cc-pVTZ level. The multireference state of 3SDs is guided by DMRG/cc-pVTZ. 34
4.4 Spin angular momentum of noisy ground states of VQE-2SDs and VQE-HF for molecules. Left: $H_2(2e, 2o)$ and Right: $F_2(10e, 6o)$. . . . 34
4.5 Angular momentum of noisy ground states of VQE-HF, VQE-3SDs, and their corresponding solutions with penalty terms for molecules. Left: $H_2O(4e, 4o)$ and Right: $N_2(6e, 6o)$. The Hamiltonian with penalty terms is denoted as $H' = H + \lambda S^2$, where $\lambda = 0.1$ for H_2O and $\lambda = 0.5$ for N_2. The λS^2 is introduced to enforce proper constraints on the quantum states by penalizing deviations from the desired values of angular momentum. 35

A.1 Convergence iterations of VQE algorithm under the IMFIL optimizer for four example systems. Left: $H_2O(4e, 4o)$ at R = 2.45 Å (top) and R = 1.85 Å(bottom), and Right: $F_2(10e, 6o)$ at R = 2.7 Å (top) and R = 2.1 Å (bottom). II

List of Tables

3.1 Statistics of Qubit Counts (N_{qb}), Layers Numbers (N_L), and Single (N_1) and Two-Qubit Gates (N_2) in the R_Y-Linear Ansatz and MR Preparation Circuits for Different Molecules. For N_1 in MR Circuit: The Numbers in Parentheses Represent the Gates Used in HF State Preparation. 29

1
Introduction

Quantum chemistry applies the principles and equations of quantum mechanics to predict and explain the reactivity, chemical, and physical properties of atoms and molecules. This predictive capability is of significant importance ranging from the design and discovery of new drugs to the understanding of the origin of life in the universe. Quantum mechanics has offered superior qualitative theoretical explanations for molecular structures, bonding, and reaction pathways, among others. However, it presents higher complexity in terms of the computational effort required for analytic and quantitative results. As Dirac claimed [1],

> *"The underlying physical laws necessary for the mathematical theory of a large part of physics and the whole of chemistry are thus completely known, and the difficulty is only that the exact application of these laws leads to equations much too complicated to be solvable."*

The inherent challenge in quantum chemistry lies in the inability to exactly solve the complex multi-particle Schrödinger equation. Only a few simple systems, such as the hydrogen atom, we can obtain analytical solutions. For many-electron systems, a series of reasonable approximations and the truncation of the Hilbert space are employed to relax the computational complexity, allowing access to effective information on electronic structure through numerical methods. The complexity of the solutions grows exponentially with the number of particles, rendering even supercomputers incapable of efficiently simulating realistic quantum systems with high precision. With the strong demand for computational power in modern computational science, the progress of supercomputers has to some extent surpassed what Dirac envisioned. However, with the size of transistors or other devices potentially reaching the scale of atomic diameters, the pursuit of classical computing power may be nearing its physical constraints. More powerful computing devices are required to execute tasks that are impossible or impractical for traditional computers. Fortunately, nature itself provides us with inspiration.

1.1 The age of quantum information

Since Feynman proposed the concept of constructing a quantum mechanical computer to efficiently handle quantum systems in 1982 [2], and Shor introduced the quantum factorization algorithm in 1994 [3], enthusiasm for exploring potential applications of quantum computing is reaching a revolutionary peak. The idea of

utilizing quantum information processing and quantum entanglement of many-body systems has evolved from cryptography [4] and communication [5] to encompass various subjects such as sensing [6], and simulation in quantum field theory [7, 8] and quantum chemistry [9, 10, 11]. In physics and chemistry, a quantum computer typically refers to a universal digital quantum simulator. It can be programmed to efficiently simulate quantum many-body systems, thus theoretically enabling precise investigation of their properties. Theoretical constructs such as non-Abelian anyon braiding [12] and discrete time crystals [13], which were once purely theoretical, can now be realized and investigated through quantum simulation.

Generalized quantum information theory (QIT) offers numerous novel perspectives and theoretical tools for understanding nature. For instance, in quantum chemistry, QIT provides a more concise definition of electronic correlation and classifies it into orbital correlation and particle correlation [14]. In quantum matter, beyond the Landau symmetry-breaking theory, the quantum phases of matter can be categorized into short-range entangled states and long-range entangled states [15]. Furthermore, the geometric structure of spacetime is regarded as an emergent feature generated by underlying quantum entanglement. The mechanism of emergence bears resemblance to quantum error-correcting codes, giving rise to research known as holographic quantum error-correcting codes [16, 17]. Recently, this toy model in the context of quantum gravity has also attracted some practical research in quantum computing [18, 19].

The fundamental unit of quantum information is the quantum bit (qubit). Unlike classical computing, which relies on Boolean variables with only two states, 0 and 1, the presence of quantum superposition allows qubits to exist in combinations of both states simultaneously. As a result, the vector space of an n-qubit system is 2^n dimensional, providing exponential potential for information processing. Engineers and experimental scientists face significant challenges in utilizing platforms such as superconductors, ion traps, nuclear magnetic resonance, neutral atoms, quantum dots, and nitrogen-vacancy centers in diamond to fabricate qubits [20, 21]. One of the major challenges is the control or mitigation of quantum decoherence. The process by which a quantum system loses its quantum properties and transitions into a classical state is called decoherence. Decoherence implies that information from the system flows into the environment. In quantum computing, any noise signals from the external environment that couple with the quantum computer can ultimately lead to erroneous outputs during computation.

In the quantum circuit model, computation entails a sequence of operations including quantum gates, measurements, initialization of qubits, and others. There are other models of quantum computation, including measurement-based quantum computation [22], adiabatic quantum computation [23], discrete-time quantum walk [24] and topological quantum computation [25]. We will refrain from exploring additional models here, as such discussions lie outside the scope of this thesis. In the quantum circuit model, each elementary operation (gate) is implemented by modifying the physical environment, such as through laser pulses or microwave radiation. During

this process, inevitable energy dissipation occurs, signifying the irreversible loss of quantum information into the environment. Therefore, the error rate of quantum gates serves as a crucial indicator for evaluating the quality of quantum computers. Concurrently, quantum dissipation on quantum devices also drives the development of the thermodynamics of quantum information [26, 27].

The theoretical framework of quantum error correction (QEC) offers vast prospects for overcoming hardware defects. The key of QEC lies the idea of encoding quantum information into a higher-dimensional Hilbert space. Quantum information can be preserved by storing it non-locally within a highly entangled state. When the environment interacts locally with parts of the system, it can only acquire minimal information about the encoded quantum state, thus avoiding disruption of the quantum state. The most favored error correction code scheme currently is the surface codes [28]. A milestone achievement was presented in Ref. [29], where up to 48 logical qubits were encoded using the surface code scheme based on Rydberg atom arrays, with over 200 logical gate operations performed. However, the logical qubits achieved here did not achieve strict error correction because neutral atoms, such as rubidium atoms, are prone to loss during gate operations and readout. Furthermore, Ref. [30] reported four highly reliable logical qubits on the ion trap platform, achieving error rates as low as 10^{-5} in their tested quantum circuits. Useful algorithms typically require extremely low gate error rates, usually below approximately 10^{-10} [31]. For the investigation of complex open-shell species such as the iron molybdenum cofactor (FeMoco) in biological nitrogen fixation, accurate computation of their electronic structure within active spaces necessitates over 50 logical qubits with extremely low error rates [32, 33]. It is evident that there is still a considerable distance to go before demonstrating fault-tolerant quantum computations. The serious overhead of QEC exceeds the current engineering capabilities.

The term "noisy intermediate-scale quantum (NISQ)" has been coined to describe quantum computers with tens to hundreds of qubits, which are subject to noise and lack error correction. NISQ algorithms are designed for quantum computers in the NISQ era, aiming to maximize the utilization of quantum computing capabilities on current devices. A notable characteristic of NISQ algorithms is their typical adoption of hybrid classical-quantum computing approaches, such as the quantum approximate optimization algorithm (QAOA) [34] and the variational quantum eigensolver (VQE) [35]. In hybrid algorithms, the quantum part typically performs specific quantum operations such as quantum state preparation, quantum gate operations, and quantum measurements. These operations aim to exploit the parallelism and entanglement inherent in quantum computing to address particular aspects of a problem. On the other hand, the role of classical part is to process and optimize the outcomes from computations, such as refining quantum parameters using classical optimization algorithms. NISQ algorithms have found many applications in quantum information scrambling, ground state and excited state preparation, machine learning, and combinatorial optimization [36]. Many of these applications often rely on error mitigation techniques. The aim of quantum error mitigation is to suppress errors through post-processing methods and multiple runs of quantum circuits,

rather than eliminating errors altogether. Quantum error mitigation methods are the focal point of this work and will be further explored in Sec. 2.3.

1.2 Quantum chemistry on quantum computers

Quantum chemistry is widely recognized as one of the earliest beneficiaries of quantum computing, poised to potentially yield a more profound and enduring impact on the field than even the well-known Shor's algorithm [37]. Aspuru-Guzik et al. were the first to demonstrate how molecular energies can be obtained on a quantum computer using quantum phase estimation (QPE) [38]. The algorithm begins with an initial reference state, which undergoes multiple controlled-unitary operations. During this process, the phase information of each eigenstate in the superposition state is amplified and encoded into the quantum state. QPE can retrieve information about the phase, thus obtaining eigenvalues, by sampling from the distribution induced by the initial state. Subsequent work has made advancements in improving QPE algorithms in terms of parallelization, efficiency, scaling, and noise resilience [9]. However, the circuit depth required for QPE is still too deep for current devices, although there have been some attempts to implement QPE on NISQ devices [39, 40]. Additionally, if the overlap between the initial state and the target eigenstate is not sufficiently large, the phase information cannot be effectively encoded, leading to algorithm failure. Despite the polynomial scaling of the QPE algorithm itself, which potentially offers exponential quantum advantage, the initial state preparation is still a non-trivial problem. Whether the initial state preparation and the overall process of QPE can provide exponential quantum advantage for quantum chemistry, even in the context of fault-tolerant quantum computing, remains an open question [41].

The variational quantum eigensolver (VQE) was designed as a NISQ algorithm to find the ground state of a given physical or chemical system [35], and has since been extensively researched. VQE reduces circuit depth but increases the overhead of polynomially scaling repeated measurements and parameter optimization. For more details on the workflow of VQE, see Sec. 2.2. In addition, the methods including orthogonal constrained VQE [42], quantum subspace methods [43], and quantum equation of motion [44] have been proposed for the calculations of low-lying excited electronic states. Another alternative method to VQE is quantum imaginary time evolution (QITE). QITE can determine not only the eigenstates of the Hamiltonian but also thermal states [45]. Furthermore, it can be extended to simulate open quantum systems [46]. Methods for obtaining more compact wavefunction representations using non-unitary operators, such as Jastrow factor [47], transcorrelated Hamiltonian [48, 49], and neural network post-processing [50], have also garnered significant attention on NISQ devices. Furthermore, quantum versions of various classical theoretical frameworks, such as quantum density functional theory [51], quantum linear response theory [52], and quantum machine learning [53], are currently under development, fostering advancements across diverse domains of chemistry research.

1.3 Thesis outline

We begin in Chapter 2 with basic theoretical background on quantum chemistry, VQE, and quantum error mitigation. Following this, we describe the idea and method of multireference-state error mitigation. In Chapter 3, we present the software tools used in this study. In Chapter 4, we demonstrate the advantages of multireference-state error mitigation through digital simulation. Finally, Chapter 5 concludes with a summary and outlook.

2

Theory

In this chapter, we review the electronic structure problem in quantum chemistry, the VQE algorithm, and the common quantum error mitigation methods. We then provide a detailed overview of the concept of reference-state error mitigation and its extension to multireference scenarios.

2.1 The electronic structure problem

2.1.1 The molecular Hamiltonian

The molecular Hamiltonian contains the information necessary to describe a molecular system. By solving the eigenvalue problem of the Hamiltonian, one can obtain the energy eigenstates and eigenvalues of the system, thereby revealing important information about the stable configurations, electronic structure, and dynamic behavior of the molecule. The molecular Hamiltonian is defined as [9]

$$\begin{aligned} H = &- \sum_i \frac{\hbar^2}{2m_e} \nabla_i^2 - \sum_I \frac{\hbar^2}{2M_I} \nabla_I^2 - \sum_{i,I} \frac{e^2}{4\pi\epsilon_0} \frac{Z_I}{|\mathbf{r}_i - \mathbf{R}_I|} \\ &+ \frac{1}{2} \sum_{i \neq j} \frac{e^2}{4\pi\epsilon_0} \frac{1}{|\mathbf{r}_i - \mathbf{r}_j|} + \frac{1}{2} \sum_{I \neq J} \frac{e^2}{4\pi\epsilon_0} \frac{Z_I Z_J}{|\mathbf{R}_I - \mathbf{R}_J|}, \end{aligned} \tag{2.1}$$

where $\mathbf{R}$ and $\mathbf{r}$ represent the coordinate vectors of the nucleus and electron, respectively. The initial two terms represent the kinetic energy of the electrons and nuclei, respectively. The subsequent three terms denote the Coulomb repulsion between the electrons and nuclei, among the electrons, and among the nuclei, respectively.

To simplify this problem, the Born-Oppenheimer approximation [54] is often employed, decoupling the nuclear and electronic degrees of freedom based on the difference in time scales of the electronic and nuclear motion. In essence, this approximation considers the electrons moving within a field of fixed nuclei. Consequently, the kinetic energy term of the nuclei can be disregarded, and the repulsion between nuclei is assumed to be constant. Under these assumptions, the electronic Hamiltonian is obtained:

$$H_e = - \sum_i \frac{\nabla_i^2}{2} - \sum_{i,I} \frac{Z_I}{|\mathbf{r}_i - \mathbf{R}_I|} + \frac{1}{2} \sum_{i \neq j} \frac{1}{|\mathbf{r}_i - \mathbf{r}_j|}, \tag{2.2}$$

where we have omitted the constant terms and adopted atomic units for conciseness [9]. The Born-Oppenheimer approximation is commonly used as a standard

reference, but it can also fail severely, especially when potential energy surfaces associated with different electronic states approach each other and display conical intersections [55].

The information of an isolated quantum system is encapsulated by the wavefunction, as described by the Schrödinger equation. Solving the non-relativistic time-independent Schrödinger equation, i.e.,

$$H\psi = E\psi \tag{2.3}$$

is one of the fundamental goals of quantum chemistry. In the Born-Oppenheimer approximation, the wavefunction can be factorized as a product of the electronic and nuclear parts

$$\psi(\mathbf{R}, \mathbf{r}) = \psi_{\mathrm{e}}(\mathbf{R}, \mathbf{r})\psi_{\mathrm{n}}(\mathbf{R}). \tag{2.4}$$

Hence, we are interested in dealing with the electronic Schrödinger equation

$$H_e\psi_e = E_e\psi_e. \tag{2.5}$$

2.1.2 The Fermionic wavefunction

The primary task in constructing many-body wave functions for fermions arises from the Pauli exclusion principle. Particles with half-integer spins, such as electrons, obey Fermi-Dirac statistics, requiring the many-body wavefunction to be antisymmetric under particle exchange:

$$\psi(\cdots, \mathbf{r}_i, \cdots, \mathbf{r}_j, \cdots) = -\psi(\cdots, \mathbf{r}_j, \cdots, \mathbf{r}_i, \cdots), \quad \forall\ i \neq j. \tag{2.6}$$

Whenever an electron is added to or removed from a many-electron system, the wavefunction must be appropriately antisymmetric to fulfill the symmetry constraints of system. This antisymmetry constraint can be described in two ways: first quantization and second quantization.

The terms "first quantization" and "second quantization" originated from historical factors during the early stages of quantum mechanics. Initially, the focus was primarily on describing individual particles, and physicists relied to some extent on classical theory as a guide. In first quantization, physical quantities such as position and momentum are quantized into operators, meaning their values are no longer definite as in classical physics, but correspond to operators in the Hilbert space representing observable quantities. In first quantization, there are fixed and finite degrees of freedom, with an emphasis on computing wavefunctions. In contrast, second quantization aims to address systems with an arbitrary number of degrees of freedom. The focus is shifted to computing the state of the system, representing it in terms of the occupation numbers of particles in different states, and calculating the probabilities of the system being in various particle occupation states.

2.1.2.1 First quantization

If we consider an N-electron wavefunction and simply take it as the product of the orthogonal spin orbitals of each electron, it is referred to as a Hartree product:

$$\psi(\mathbf{x}_1, \mathbf{x}_2, \cdots, \mathbf{x}_N) = \chi_1(\mathbf{x}_1)\chi_2(\mathbf{x}_2)\cdots\chi_N(\mathbf{x}_N), \tag{2.7}$$

where $\mathbf{x_i} = \{\mathbf{r_i}, \sigma_i\}$ denotes the spatial and spin coordinate of an electron. This construction does not adhere to the antisymmetry principle necessary for fermionic systems. The Slater determinant is introduced to solve this problem [56]

$$\psi(\mathbf{r}_1, \mathbf{r}_2, \ldots, \mathbf{r}_N) = \frac{1}{\sqrt{N!}} \begin{vmatrix} \chi_1(\mathbf{x}_1) & \chi_2(\mathbf{x}_1) & \cdots & \chi_N(\mathbf{x}_1) \\ \chi_1(\mathbf{x}_2) & \chi_2(\mathbf{x}_2) & \cdots & \chi_N(\mathbf{x}_2) \\ \vdots & \vdots & \ddots & \vdots \\ \chi_1(\mathbf{x}_N) & \chi_2(\mathbf{x}_N) & \cdots & \chi_N(\mathbf{x}_N) \end{vmatrix} \tag{2.8}$$

which is an antisymmetrized product of the single-electron basis functions.

In a Slater determinant, the columns represent different spin orbitals and the rows represent different electrons. Interchanging the coordinates of two electrons corresponds to swapping the two rows of the Slater determinant. This operation results in a change of sign for the determinant. Additionally, if two electrons occupy the same orbital, then the determinant evaluates to zero. Thus, a Slater determinant underscores the antisymmetry principle and the Pauli exclusion principle. However, the utilization of first quantization remains redundant for indistinguishable particles, as we still explicitly represent many-particle wavefunctions using the coordinates of individual electrons.

2.1.2.2 Second quantization

Another approach to managing the antisymmetry condition is through second quantization. In second quantization, we use Fock space to describe the states of multi-particle systems. Fock space is a space constructed as a direct sum of single-particle states, encompassing all possible combinations of particle numbers and their states. Multi-particle states in Fock space can be easily manipulated by introducing creation and annihilation operators. Creation operators add a particle to the system, while annihilation operators remove a particle from the system, i.e.,

$$\begin{aligned} a_\alpha^\dagger|\cdots, n_\beta, n_\alpha, n_\gamma, \cdots\rangle &= (-1)^{\sum_{\beta<\alpha} n_\beta}\sqrt{1-n_\alpha}|\cdots, n_\beta, 1+n_\alpha, n_\gamma, \cdots\rangle, \\ a_\alpha|\cdots, n_\beta, n_\alpha, n_\gamma, \cdots\rangle &= (-1)^{\sum_{\beta<\alpha} n_\beta}\sqrt{n_\alpha}|\cdots, n_\beta, 1-n_\alpha, n_\gamma, \cdots\rangle. \end{aligned} \tag{2.9}$$

where $|n\rangle \equiv |\cdots, n_\beta, n_\alpha, n_\gamma, \cdots\rangle$ is known as a Fock state or occupation number state. From a vacuum state $|0\rangle$, we can construct arbitrary SDs (SDs) by applying a product of creation operators. Creation and annihilation operators are constructed based on the anti-commutation property of fermions:

$$\begin{aligned} \{a_p, a_q^\dagger\} &= a_p a_q^\dagger + a_q^\dagger a_p = \delta_{pq}, \\ \{a_p, a_q\} &= \{a_p^\dagger, a_q^\dagger\} = 0, \end{aligned} \tag{2.10}$$

thus it enforces the antisymmetry principle.

In second quantization, the electronic Hamiltonian Eq. 2.2 can be expressed as

$$H = \sum_{p,q} h_{pq} a_p^\dagger a_q + \frac{1}{2} \sum_{p,q,r,s} h_{pqrs} a_p^\dagger a_q^\dagger a_r a_s, \tag{2.11}$$

with

$$\begin{aligned} h_{pq} &= \int d\mathbf{x} \phi_p^*(\mathbf{x}) \left(-\frac{\nabla^2}{2} - \sum_I \frac{Z_I}{|\mathbf{r} - \mathbf{R}_I|} \right) \phi_q(\mathbf{x}), \\ h_{pqrs} &= \int d\mathbf{x}_1 d\mathbf{x}_2 \frac{\phi_p^*(\mathbf{x}_1) \phi_q^*(\mathbf{x}_2) \phi_r(\mathbf{x}_2) \phi_s(\mathbf{x}_1)}{|\mathbf{r}_1 - \mathbf{r}_2|}. \end{aligned} \tag{2.12}$$

The first one-electron integral represents the kinetic energy of the electrons and their Coulomb interaction with the nuclei. The second two-electron integral arises from the Coulomb repulsion between electrons. $\phi_p(r) \equiv \langle r|\phi_p\rangle$ is known as an atomic orbital basis function.

The collection of these basis functions constitutes a basis set. Linear combinations of atomic orbitals construct molecular orbitals. Gaussian-type orbitals (GTOs) are the most commonly used orbitals to date. They offer a computationally convenient formulation for evaluating the integrals mentioned above. Basis sets vary in the number and complexity of included basis functions, with larger basis sets generally offering more accurate results at the expense of higher computational costs. The selection of a basis set significantly impacts computational results as it determines the form and approximation level of the wavefunction. For further exploration of basis sets, refer to [57, 58].

2.1.3 Electron correlation

The treatment of electron correlation plays an indispensable role in accurate quantum chemical calculations. In the Hartree-Fock (HF) method, each electron is treated as moving in the mean field generated by the other electrons, neglecting their specific interactions. This effectively reduces the N-electron problem to N individual electron problems, without considering electron correlation. The HF ground state in second quantization can be expressed as

$$|\psi_0\rangle = a_N^\dagger \dots a_2^\dagger a_1^\dagger |0\rangle^{\otimes N}. \tag{2.13}$$

The HF method aims to find the dominant Slater determinant as an approximation to the electronic wavefunction. However, failures of this single determinant model are common, such as in open-shell molecules, excited states and transition metal compounds. The complete electronic wavefunction should be expressed as a linear combination of SDs:

$$|\psi\rangle = \sum_I c_I |\psi_I\rangle = c_0|\psi_0\rangle + \sum_{I\geq 1} c_I |\psi_I\rangle . \tag{2.14}$$

Considering all possible electronic configurations leads to exact energy calculations for a given basis set, known as full configuration interaction (FCI). The numerically

exact solution of the time-independent non-relativistic Schrödinger equation is obtained when a finite basis set is expanded to a complete basis set. The number of determinants required for a FCI expansion grows factorially with the number of electrons and orbitals [59]:

$$N_{\text{det}} = \binom{N}{N_\alpha}\binom{N}{N_\beta}, \tag{2.15}$$

with N spatial orbitals and N_α/N_β the number of alpha/beta electrons. The factorial scaling currently limits FCI to systems with at most around 20 electrons in practice [60].

In general, the difference between the HF energy and the exact ground state energy (FCI) is defined as (basis set) electron correlation energy, i.e.,

$$E_{\text{corr}}^{\text{basis}} = E_{\text{exact}}^{\text{basis}} - E_{\text{HF}}^{\text{basis}}. \tag{2.16}$$

Correlation arises from the instantaneous interaction between electrons, which can be divided into dynamic and static correlations. Dynamic correlation is typically regarded to be short-range. It is predominantly governed by the Hartree-Fock determinant, with minor supplementary contributions from potentially excited configurations. This can be effectively captured by single-reference correlation methods, such as configuration interaction singles and doubles (CISD) and coupled-cluster singles and doubles (CCSD). Static correlation, on the other hand, is considered long-range and arises, for instance, during bond breaking, where multiple near-degenerate SDs exist. This necessitates the use of multi-reference correlation methods, such as multi-configuration self-consistent field (MCSCF) and multi-reference perturbation theory (MRPT).

2.1.4 Active space

Calculations based on the degrees of freedom of all orbital occupations in FCI incur prohibitive costs. To address this, the complete active space (CAS) approach involves selecting a subset of molecular orbitals most relevant to the electronic states of interest (referred to as the active space) and including only the active electrons that are allowed to occupy these active orbitals. The complete orbital space can be divided into inactive orbitals, active orbitals, and virtual orbitals, as shown in Fig. 2.1. The complete active space SCF method (CASSCF) is a special case of MCSCF within the CAS, where both the molecular orbitals and CI coefficients are optimized. The CASSCF method encompasses all possible electronic configurations within a given CAS. It performs FCI calculations on these configurations, thereby capturing significant correlation effects. Notably, orbital rotations between active-virtual and active-inactive orbitals contribute to including some effects of out-of-CAS orbitals. This approach balances computational efficiency while capturing strong correlations within the activity space. However, the cost of CASSCF still scales exponentially with the growth of the active space. Some promising multi-reference methods have been proposed to approximate the CASSCF solver, such as density matrix renormalization group (DMRG) [61], full configuration interaction quantum Monte Carlo (FCIQMC) [62], selected configuration interaction (SCI) [63],

and density matrix embedding theory (DMET) [64].

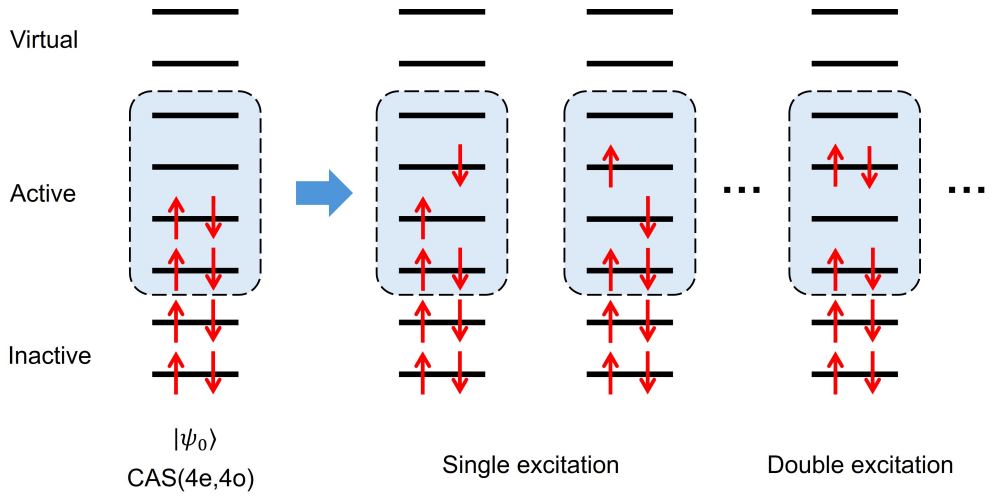


Figure 2.1: Determinants in a CAS(4e,4o) complete active space, where there are 4 active electrons in 4 active orbitals, and examples of single and double excitations.

2.2 Variational quantum eigensolver

2.2.1 Overview of VQE

The variational quantum eigensolver (VQE) is a promising algorithm for near-term quantum computers. VQE is designed to find an approximate ground state of the Hamiltonian based on the Rayleigh-Ritz variational principle,

$$E_0 \leq \langle \Psi(\boldsymbol{\theta}) | H | \Psi(\boldsymbol{\theta}) \rangle, \tag{2.17}$$

where E_0 is the ground state energy of the Hamiltonian H. VQE relies on a parameterized quantum circuit (PQC), $U(\boldsymbol{\theta})$, to prepare an ansätze quantum state, $|\Psi(\boldsymbol{\theta})\rangle = U(\boldsymbol{\theta})\,|\Psi_0\rangle$, and accesses the ground state of the Hamiltonian by iteratively optimizing the parameters $\boldsymbol{\theta}$ classically. The state $|\Psi_0\rangle$ is an initial reference state that has a large overlap with the the ground state and is expected to be efficiently prepared.

The workflow of VQE is illustrated in Fig. 2.2; it includes [65]:

1. Map the problem Hamiltonian to quantum hardware in the form of Pauli strings P_α,
$$H = \sum_\alpha h_\alpha P_\alpha = \sum_\alpha h_\alpha \bigotimes_{j=1}^{N} \sigma_j^{\alpha_j}. \tag{2.18}$$
2. Prepare an ansätze circuit $U(\boldsymbol{\theta})$ with a set of parameters $\boldsymbol{\theta}$ to generate a trial state $|\Psi(\boldsymbol{\theta})\rangle$.
3. Sample and repeatedly measure each Pauli string under the same trial state, determining the expectation value of each Pauli string, $\langle \Psi(\boldsymbol{\theta}) | P_\alpha | \Psi(\boldsymbol{\theta}) \rangle$.

4. Calculate the cost function

$$E(\boldsymbol{\theta}) = \sum_{\alpha} h_\alpha \langle \Psi(\boldsymbol{\theta}) | P_\alpha | \Psi(\boldsymbol{\theta}) \rangle \tag{2.19}$$

by summing up the results of the measurements.
5. Optimize $E(\boldsymbol{\theta})$ classically and update the parameters. Repeat the previous steps until convergence criteria.

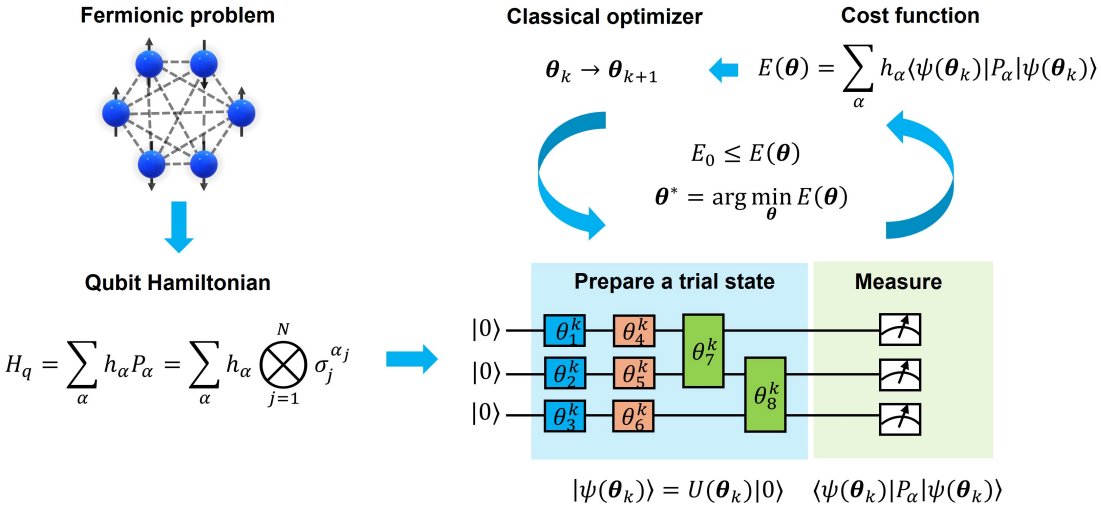


Figure 2.2: Graphical representation of the workflow of VQE.

2.2.2 Mapping to the qubit space

The first step of electronic structure simulation on quantum computers requires encoding fermionic systems onto qubits. Algorithms can encode the fermionic Hamiltonian using either first or second quantization. First quantization merges antisymmetry into the qubit basis itself, while second quantization encodes antisymmetry into qubit operators rather than quantum states, providing advantages in qubit reduction [66, 67]. Therefore, we will focus on second quantized mapping methods.

2.2.2.1 The Jordan-Wigner mapping

The most direct fermion-to-qubit mapping method is the Jordan-Wigner (JW) mapping. It directly maps the occupation of a spin orbital to the occupation of a qubit,

$$|f_{N-1}, ..., f_1, f_0\rangle \rightarrow |q_{N-1}, ..., q_1, q_0\rangle, \quad q_j = f_j \in \{0, 1\}. \tag{2.20}$$

The fermionic creation and annihilation operators are given by

$$\begin{aligned} a_j^\dagger &\rightarrow Q_j^\dagger \otimes Z_{j-1} \otimes \cdots \otimes Z_0, \\ a_j &\rightarrow Q_j \otimes Z_{j-1} \otimes \cdots \otimes Z_0, \end{aligned} \tag{2.21}$$

where $Q_j^\dagger = (X_j - iY_j)/2$ and $Q_j = (X_j + iY_j)/2$ are the qubit creation and annihilation operators in terms of Pauli matrices $X_j = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}$ and $Y_j = \begin{bmatrix} 0 & -i \\ i & 0 \end{bmatrix}$.

The Pauli weight of the JW mapping scales as $\mathcal{O}(N)$ due to the direct mapping. The Pauli weight indicates the maximum count of non-identity operators within any Pauli string generated by the mapping. Encoding with low Pauli weight can lead to circuit construction achieving target accuracy with lower depth and fewer entangling gates.

2.2.2.2 The parity mapping

An alternative to the JW mapping is parity mapping. Instead of storing occupancy information locally, the parity encoding method encodes the parity of all occupied orbitals up to j-th orbital locally on the qubit j [68]. In the parity mapping, one represents $|0\rangle_j$ if the number of occupied orbitals up to the j-th orbital is even, and $|1\rangle_j$ if it is odd. In the parity mapping, the transformation of a fermionic state involves

$$| f^{\beta}_{n_\beta-1}..., f^{\beta}_0, f^{\alpha}_{n_\alpha-1}, ..., f^{\alpha}_0\rangle \rightarrow \left|p_{n_\alpha+n_\beta-2}, ..., p_{n_\alpha-1}, ..., p_1, p_0\right\rangle \tag{2.22}$$

where the ordering the fermionic state is called spin-block order and

$$p_i = \sum_j [\pi_n]_{ij} f_j \pmod 2, \quad [\pi_n]_{ij} = \begin{cases} 1 & i < j \\ 0 & i \geq j \end{cases}, \tag{2.23}$$

so that π_n denotes a $n \times n$ matrix

$$\pi_n = \begin{pmatrix} 1 & 1 & \cdots & 1 \\ 0 & 1 & \cdots & 1 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & 1 \end{pmatrix}. \tag{2.24}$$

The occupancy of the orbital j can be determined based on the change in parity at the position j. If the parity changes, i.e., $|01\rangle_{j-1,j}$ or $|10\rangle_{j-1,j}$, the orbital j is occupied. Conversely, if the parity does not change, i.e., $|00\rangle_{j-1,j}$ or $|11\rangle_{j-1,j}$, the orbital j is unoccupied. The transformed creation and annihilation operators are given by [69]

$$\begin{aligned} a^\dagger_j &\rightarrow \frac{Z_{j-1} \otimes X_j - iY_j}{2} \otimes X_{j+1} \otimes \cdots \otimes X_{n-1} \\ a_j &\rightarrow \frac{Z_{j-1} \otimes X_j + iY_j}{2} \otimes X_{j+1} \otimes \cdots \otimes X_{n-1} \end{aligned} \tag{2.25}$$

Additionally, we know that

$$p_{n_\alpha} = \sum_{j=1}^{n_\alpha} f^\alpha_j = N_\alpha, \quad p_{n_\alpha+n_\beta} = \sum_{j=1}^{n_\alpha} f^\alpha_j + \sum_{j=1}^{n_\beta} f^\beta_j = N_\alpha + N_\beta, \tag{2.26}$$

implying an additional two-qubit reduction for spin-block ordering in the parity mapping. This reduction stems from the known values of these two qubits in a system that conserves the number of alpha and beta electrons. The Pauli weight of the parity mapping scales as $\mathcal{O}(N)$, similar to the JW mapping.

2.2.2.3 The Bravyi-Kitaev mapping

The Bravyi-Kitaev (BK) mapping encodes locally a combination of parity and occupation number in qubits, based on the BK matrix β_n [68],

$$|f_{N-1}, ..., f_1, f_0\rangle \rightarrow |b_{N-1}, ..., b_1, b_0\rangle, \quad b_i = \sum_j [\beta_n]_{ij} f_j \pmod 2, \tag{2.27}$$

where β_n is recursively defined by

$$\beta_1 = [1], \quad \beta_{2^{i+1}} = \begin{pmatrix} \beta_{2^i} & \mathbf{A} \\ \mathbf{0} & \beta_{2^i} \end{pmatrix}, \tag{2.28}$$

and the matrix $\mathbf{A}$ has its top row filled with ones, while the rest of its entries are zeros. For example, employing β_8 results in the following transformation (all sums in mod 2):

$$\begin{pmatrix} 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 \\ 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 1 & 1 & 1 \\ 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 & 1 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} 1 \\ 0 \\ 1 \\ 0 \\ 0 \\ 1 \\ 1 \\ 1 \end{pmatrix} = \begin{pmatrix} 1 \\ 0 \\ 1 \\ 0 \\ 1 \\ 1 \\ 0 \\ 1 \end{pmatrix}. \tag{2.29}$$

The BK mapping exhibits better scaling of Pauli weight as $\mathcal{O}(\log_2(n))$, but it demonstrates weaker resilience to noise [70].

2.2.2.4 Qubit tapering

The qubit tapering techniques exploit the symmetry inherent in the molecular Hamiltonian to reduce the number of qubits needed in quantum computing. This approach is beneficial for both quantum resource and noise reduction in VQE. In the different encodings, some qubits represent conserved quantities of the molecular system and they can be targeted for qubit reduction. An example is the two-qubit reduction mentioned earlier in the context of parity mapping. If all Pauli strings in the qubit Hamiltonian commute with an operator, then that operator is considered a symmetry operator. For example, particle number operators for α and β electrons fulfil

$$[H, N_\alpha] = [H, N_\beta] = 0. \tag{2.30}$$

The main idea behind symmetry tapering is to find a unitary U that transforms the original Hamiltonian into a new Hamiltonian with the same eigenvalues [71, 72],

$$H' = U^\dagger H U = \sum_\alpha h_\alpha \Gamma_\alpha, \quad \Gamma_\alpha = U^\dagger P_\alpha U. \tag{2.31}$$

We expect that all terms in H' commute with all Pauli-X operators on a set of qubits $\{j\}$

$$[H', X_{\{j\}}] = 0 \tag{2.32}$$

This is feasible only if each Γ_α in H' acts trivially on a set of qubits $\{j\}$ by operators I or X [71]. This means that the qubits $\{j\}$ can be eliminated by replacing the Pauli-X operators $X_{\{j\}}$ with their eigenvalues ± 1.

The unitary operator U is constructed based on the $\mathbb{Z}_2$ symmetries. For each specific symmetry τ_i, it is possible to eliminate one qubit in the qubit Hamiltonian. The unitary satisfies

$$U\tau_i U^\dagger = X_{q(i)}, \tag{2.33}$$

where the symmetry operator acts as Pauli-X operators on the qubit $q(i)$ after transformation. The transformation U is given by $U = \prod_i U_i$ with

$$U_i = \frac{X_{q(i)} + \tau_i}{\sqrt{2}}. \tag{2.34}$$

It is worth noting that the qubit tapering reported in Ref [71] exploits electron number conservation and spin symmetry. Consequently, this qubit tapering technique preserves the ground state of the system while modifying the accessible excited states.

2.2.3 Parameterized ansätze

Designing and selecting the appropriate ansätze is crucial in VQE. The design of ansätze is heuristic, drawing inspiration from traditional quantum chemistry methods, practical hardware limitations, and specific physical principles and constraints. This approach prioritizes practical experience and intuitive adaptation over rigorous theoretical derivations. From a structural perspective, parameterized ansätze can be divided into fixed and adaptive structures. An adaptive ansätze selectively adds operators that contribute most to lowring the energy during each iteration, resulting in a more compact circuit. However, reference state error mitigation (see Sec. 2.3.5) requires a fixed ansätze structure to ensure consistency in gate noise. Fixed structure ansätze include chemistry-inspired, hardware-efficient, symmetry-preserving, and Hamiltonian variational ansätze. We will focus on the first three types.

2.2.3.1 Chemistry-inspired ansätze

One type of ansätze relies on established principles that guide the formulation of trial wavefunctions. These principles are rooted in conventional theories of quantum chemistry, providing a structured approach to designing trial wavefunctions that reflect fundamental concepts and interactions within molecular systems. An example is the unitary coupled cluster (UCC) form. The UCC ansätze facilitates the generation of excitations from the initial reference state by employing the exponentiated excitation operator:

$$U(\boldsymbol{\theta}) = e^{T - T^\dagger}. \tag{2.35}$$

Here, $T = \sum_i \theta_i T_i$ is typically truncated at excitation level $i = 2$, incorporating single and double excitation operators:

$$T_{UCCSD} = T_1 + T_2 = \sum_{i \in virt, \alpha \in occ} t_i^\alpha a_i^\dagger a_\alpha + \sum_{i,j \in virt, \alpha,\beta \in occ} t_{ij}^{\alpha\beta} a_i^\dagger a_j^\dagger a_\beta a_\alpha. \tag{2.36}$$

The unitary operator $U(\boldsymbol{\theta})$ needs to be decomposed into operations implementable on a quantum device, often achieved through Trotter-Suzuki decomposition:

$$\begin{aligned} e^{T-T^\dagger} &= e^{\sum_i \theta_i (T_i - T_i^\dagger)} \\ &\approx \left(\prod_i e^{\frac{\theta_i}{t}(T_i - T_i^\dagger)} \right)^t + \mathcal{O}\left(\frac{1}{t}\right), \end{aligned} \tag{2.37}$$

where t is the order of decomposition (Trotter number). In VQE, a Trotter number of $t = 1$ is usually sufficient for small molecules in VQE.

The number of gates of UCCSD in the JW mapping scales as $\mathcal{O}(K^2N^2)$ for K spin orbitals and N electrons [11]. The unfavorable scaling of UCCSD has driven research into its variants. One extension is the k-unitary pair coupled-cluster generalized singles and doubles (k-UpCCGSD) [73]. Here, the term "generalized" implies that the single and double excitation terms do not differentiate between occupied and unoccupied orbitals. Furthermore, "pair coupled-cluster" indicates that the double excitations exclusively involve two-body excitations, moving pairs of opposite-spin electrons together from doubly-occupied to unoccupied spatial orbitals. k-UpCCGSD enables linear scaling of $\mathcal{O}(kN)$, where k denotes the number of repetitions of the excitation operator. This approach can attain higher accuracy at a lower cost of quantum resources. Ref. [74] provides a unified perspective on the UCC ansätze family.

2.2.3.2 Hardware-efficient ansätze

One type of ansätze is specifically crafted to harness the capabilities of available hardware, ensuring practical feasibility and efficiency in real-world computational applications. Hardware-efficient ansätze (HEA) are designed with the objective of minimizing resource requirements, such as the number of gates and connections, necessary for advancing the applicability and scalability of quantum algorithms in computational practice. HEA are typically composed of repeated blocks of parameterized rotation gates and entangling gates [75],

$$\begin{aligned} |\Psi(\boldsymbol{\theta})\rangle = \prod_{q=1}^{N} [U^{q,d}(\boldsymbol{\theta})] \times U_{\text{ent}} \times \prod_{q=1}^{N} [U^{q,d-1}(\boldsymbol{\theta})] \times \cdots \times U_{\text{ent}} \times \\ \prod_{q=1}^{N} [U^{q,0}(\boldsymbol{\theta})] |\Psi_0\rangle. \end{aligned} \tag{2.38}$$

One advantage of HEA is their low dependency on specific problem instances, enabling their application across various quantum computing scenarios. However, this broad applicability may come with the drawback of potentially not performing optimally on all tasks. The high expressivity of HEA implies a higher susceptibility to suffering from barren plateaus [76, 77], leading to an exponential increase in the resources required for training. Ref. [78] further suggested that even shallow HEA architectures should avoid tasks where the data satisfies the volume law of entanglement. One potential improvement is to employ variational entangling gates with

adjustable parameters to avoid excessive entanglement [79]. This approach can also be utilized for symmetry enhancement [80].

Both HEA and UCC ansätze can lead to non-differentiable tups in the potential energy surface (PES), stemming from the failure to preserve certain symmetries, such as electron number and total spin [81]. For example, when seeking the singlet ground state, VQE may inadvertently access triplet excited states, leading to mixing between electronic states of different symmetries (intersystem crossing). The symmetry breaking can become more pronounced under noisy environment. Therefore, additional constraints on symmetry preservation are required. One approach to address this problem is by adding a penalty term to the Hamiltonian [82], thus the cost function of VQE becomes

$$E = \langle\psi(\boldsymbol{\theta})|H|\psi(\boldsymbol{\theta})\rangle + \sum_i \lambda_i \left[\langle\psi(\boldsymbol{\theta})|S_i|\psi(\boldsymbol{\theta})\rangle - s_i\right]^2, \tag{2.39}$$

where S_i is the symmetry operator for constraints (e.g., N, S_z, S^2), λ_i is the penalty factor requiring empirical selection, s_i is the desired value of the expectation value of S_i. This penalty-based approach can also be used to access excited states.

2.2.3.3 Symmetry-preserving ansätze

Another approach to respecting symmetry is to utilize particle-conserving exchange gates to construct the ansätze, ensuring that only states with the appropriate symmetry are generated. The ASWAP ansätze is proposed in Ref. [83] by using the gate

$$A(\theta,\phi) = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & \cos\theta & e^{i\phi}\sin\theta & 0 \\ 0 & e^{-i\phi}\sin\theta & -\cos\theta & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \tag{2.40}$$

to achieve single qubit excitation. The gate A is constrained to span only the single-excitation subspace, thereby only mixing states $|01\rangle$ and $|10\rangle$. In the spin-block order (recall Eq. 2.22) of JW mapping, this arrangement ensures that the first half qubits represents the α-electron orbitals, while the remaining qubits represent β-electron orbitals. The A gates are applied on all nearest-neighbor two-qubit pairs, forming a SWAP network, except between the qubits bridging the two spin subspaces. It restricts the accessible Hilbert space, avoiding mixing of spins. A broader set of symmetries, including particle number, time reversal, total spin, and spin magnetization, can be preserved by modifying the ansätze [83]. In addition, the quantum number preserving (QNP) ansätze [84] and the tiled unitary product states (tUPS) ansätze [85] are constructed based on four-qubit spatial orbital rotation gates and four-qubit double-excitation gates, providing high-accuracy and spin preservation. The tUPS ansätze exhibits faster convergence, connections to valence bond theory and gate-efficiency. However, a key limitation of these ansätze is their effectiveness solely under the JW mapping, where the occupation of a single electron is locally stored on a qubit.

2.3 Quantum error mitigation

Quantum error mitigation (QEM) provides essential support for the practicality of NISQ algorithms. In current quantum devices, errors stem from various sources such as measurements, environmental noise, and imperfections in control operations. These errors can propagate and accumulate throughout the system via quantum entanglement. Computations would become meaningless without addressing these errors. Quantum noise can flatten the energy landscape of VQE, resulting in changes to the optimal cost function value, shifts in the positions of minima, and symmetry breaking of degenerate global minima into local minima [86, 87]. Assessing the efficacy of error mitigation methods implies determining the extent to which these issues can be corrected. While few methods exist to directly restore noise landscapes to enhance trainability, most error mitigation techniques can partially recover ideal expectation values and measurement statistics of observables. Moreover, many methods can be combined simultaneously.

2.3.1 Measurement error mitigation

Measurement readout error stands as a significant source of error in current quantum devices, particularly for algorithms like VQE that involves a large measurement overhead. Typical methods for measurement error mitigation rely on quantum detector tomography (QDT) [88]. QDT is uesd to generate a confusion matrix T for a specific device. The matrix T is defined by $C_{\text{noisy}} = TC_{\text{ideal}}$, where C_{noisy} represents the probability distribution after noisy samplings and C_{ideal} represents the true probability distribution. Each element $T_{i,j}$ in the matrix T encodes the conditional probability of reading out $|i\rangle$ given that the true outcome should be $|j\rangle$. To mitigate errors, we require the inverse matrix T^{-1}, such that $C_{\text{ideal}} = T^{-1}C_{\text{noisy}}$. It is worth noting that the matrix T characterizes the readout error of the device, independent of the circuit. Additionally, the matrix T might be non-invertible, and applying T^{-1} could lead to issues of negative probabilities, necessitating the introduction of additional approximations or statistical tools.

Since T is a $2^n \times 2^n$ matrix for an n-qubit system, obtaining the complete confusion matrix typically necessitates $\mathcal{O}(2^n)$ calibration experiments, which is not scalable. If one approximates that the crosstalk between different qubits during the measurement process can be neglected, then T can be approximated as the tensor product of 2×2 matrices, $T \approx T_1 \otimes T_2 \cdots \otimes T_N$. This local readout mitigation method only incurs a cost of $\mathcal{O}(N)$ number of experiments. Scalable mitigation of correlated measurement errors remains an ongoing area of research [89, 90].

2.3.2 Zero noise extrapolation

Zero noise extrapolation (ZNE) involves systematically scaling noise and fitting the results obtained under different noise levels to extrapolate the noise-free expectation values. The simplest method to increase the noise level of a circuit is to intentionally increase its depth by unitary folding [91]. We can transform the unitary circuit as

$U \to U(U^\dagger U)^n$, and define a scaling factor λ_n as the ratio of layers in the folded circuit over the original circuit, given by $\lambda = 2n + 1$. Thus, the scaled noise level can be approximately defined as $\tau' = \lambda_n \tau$. The expectation value of an observable, $\langle O \rangle(\lambda_n)$, can be described as a function of the scaling factor. Then, we aim to obtain an extrapolation model $\langle \tilde{O} \rangle(\boldsymbol{\alpha}, \lambda_n)$, where $\boldsymbol{\alpha}$ represents the fitting parameters of this model. We optimize the loss function $f(\boldsymbol{\alpha})$, defined as $f(\boldsymbol{\alpha}) = \sum_n |\langle O \rangle(\lambda_n) - \langle \tilde{O} \rangle(\boldsymbol{\alpha}, \lambda_n)|^2$, to find the optimal model parameters $\boldsymbol{\alpha}^*$ [69]. Thus, we can use this model to infer $\langle \tilde{O} \rangle(\boldsymbol{\alpha}^*, \lambda_n = 0)$ as the error-mitigated expectation value. The selection of the extrapolation model can be performed using data analysis methods such as Richardson extrapolation or machine learning.

2.3.3 Probabilistic error cancellation

The objective of probabilistic error cancellation (PEC) is to reverting the effects of noise by applying the inverse of the noise channel, denoted as Λ^{-1}, before the noisy channel Λ, thus achieving the ideal gate operation. The mitigated state reads

$$\rho = \prod_i \Lambda_i^{-1} \circ \Lambda_i \circ \mathcal{U}_i(\rho_0), \tag{2.41}$$

where $\mathcal{U}_i(\rho_0) = U \rho_0 U^\dagger$. Typically, Λ^{-1} is mathematically viable but physically unfeasible. It is not a completely positive, trace preserving (CPTP) map, thus it is not permissible as an operation in quantum mechanics. However, the following decomposition is possible:

$$\Lambda^{-1} = \sum_n \alpha_n \mathcal{B}_n, \tag{2.42}$$

where α_n is the quasi-probability weight (possibly negative) and $\mathcal{B}_n$ is the operation that is implementable on a quantum computer (which can be noisy). In other words, ideal gates can be formulated as linear combinations of realizable noisy gates [92].

PEC requires a noise model assumption that effectively characterizes the noise in quantum circuits. The noise channel Λ can be modelled as a Markovian stochastic Pauli noise channel [93],

$$\Lambda(\rho) = \sum_{a=0}^{4^n-1} \gamma_a P_a \rho P_a, \tag{2.43}$$

where $P_a \in \{I, X, Y, Z\}$ and hence the coefficients γ_a have $4^n - 1$ degrees of freedom. An alternative expression is presented as [94]

$$\Lambda(\rho) = \frac{1}{2^n} \sum_{b=0}^{4^n-1} \lambda_k \operatorname{Tr}(P_b(\rho)) P_b. \tag{2.44}$$

Here,

$$\lambda_b = \sum_a \gamma_a (-1)^{\langle a,b \rangle} \tag{2.45}$$

is termed the Pauli fidelity, with $\langle a, b \rangle$ equal to 0 if P_a commutes with P_b, and 1 otherwise. The learning of Pauli noise models is inspired by randomized benchmarking [95]. This entails multiple measurements of the expectation values of a

set of Pauli operators as the number of repetitions of the circuit increases, and fitting the decay rate of the resulting curve to ascertain the Pauli fidelities $\{\lambda_b\}$ of the noise channel [96]. The parameters $\{\gamma_a\}$ can then be determined by the relation Eq. 2.45. The Pauli noise model can be effectively inverted, and its quasi-probability distribution can be accurately sampled [96, 97]. PEC is the most accurate error mitigation method, but it comes at the cost of extensive modifications to the circuit and exponential sampling overhead.

2.3.4 Symmetry verification by post-selection

As previously mentioned, in the design process of VQE or cost functions, we can impose symmetry constraints. Additionally, we can filter out sampling results with erroneous expectations by calculating the expectation values of symmetry operators (e.g., N, N_α, N_β, S_z, etc.) with respect to each measurement result per sampling. The post-selection method for symmetry verification can effectively reduce the noisy expectation values in VQE and enhance the fidelity of the noisy ground state without introducing additional quantum costs, particularly for shallow circuits. It is worth noting that this post-selection method is a purely classical post-treatment for final results, distinct from the quantum circuit approach for ancilla symmetry verification, which embeds the symmetry verification onto the qubits within the system itself [98].

2.3.5 Reference-state error mitigation

Reference-state error mitigation (REM) is a chemically inspired, cost-effective, noise type-independent method of QEM for NISQ chemistry algorithms [99]. Both fault-tolerant algorithms and NISQ algorithms for quantum chemistry generally rely on the availability of a readily preparable initial reference state that exhibits a favorable overlap with the target ground state. Under the constraint of a fixed circuit structure, there is reason to believe that quantifying the effect of noise on the preparation circuit for reference states can aid in capturing a substantial portion of the impact of noise on circuits for ground states.

The REM approach is shown as Fig. 2.3. The first and most critical step in REM is to select and prepare high quality reference states on a quantum computer. In an ideal scenario, we expect that the reference states can be prepared by a circuit composed solely of Clifford gates (Hadamard, phase gate S, and CNOT)[1] or with the addition of a small amount of non-Clifford resources (e.g., T gate, rotation gates). Clifford circuits can be efficiently simulated on classical computers owing to the Gottesman-Knill theorem [100], facilitating the efficient determination of precise values, $E_{\text{Exact}}(\boldsymbol{\theta}_{\text{Ref}})$. However, Clifford gates alone cannot achieve universal quantum computation. The addition of any non-Clifford gate can create a universal gate set, such as $\{\text{Clifford}, T\}$. Therefore, our goal should be to prepare the desired reference state using as few non-Clifford resources as possible.

[1]Each Pauli matrix can be constructed using the phase and Hadamard gates, making it an element of the Clifford group.

Subsequently, the noisy energy of the reference state, $E_{\mathrm{VQE}}(\boldsymbol{\theta}_{\mathrm{Ref}})$, is determined based on measurements from an actual quantum device. The impact of noise on the reference state is thus quantified as

$$\Delta E_{\mathrm{REM}} = E_{\mathrm{VQE}}(\boldsymbol{\theta}_{\mathrm{Ref}}) - E_{\mathrm{Exact}}(\boldsymbol{\theta}_{\mathrm{Ref}}). \tag{2.46}$$

The noisy VQE experiment yields $E_{\mathrm{VQE}}(\boldsymbol{\theta}_{\mathrm{min,VQE}})$, and ΔE_{REM} can be employed to provide an error-mitigated result

$$E_{\mathrm{REM}}(\boldsymbol{\theta}_{\mathrm{min,VQE}}) = E_{\mathrm{VQE}}(\boldsymbol{\theta}_{\mathrm{min,VQE}}) - \Delta E_{\mathrm{REM}}. \tag{2.47}$$

In principle, the exact energy for any parameter can be represented as

$$E_{\mathrm{Exact}}(\boldsymbol{\theta}) = E_{\mathrm{REM}}(\boldsymbol{\theta}) - \Delta E_p(\boldsymbol{\theta}), \tag{2.48}$$

where $\Delta E_p(\boldsymbol{\theta})$ represents the dependence of noise on parameters and $\Delta E_p(\boldsymbol{\theta}_{\mathrm{Ref}}) = 0$.

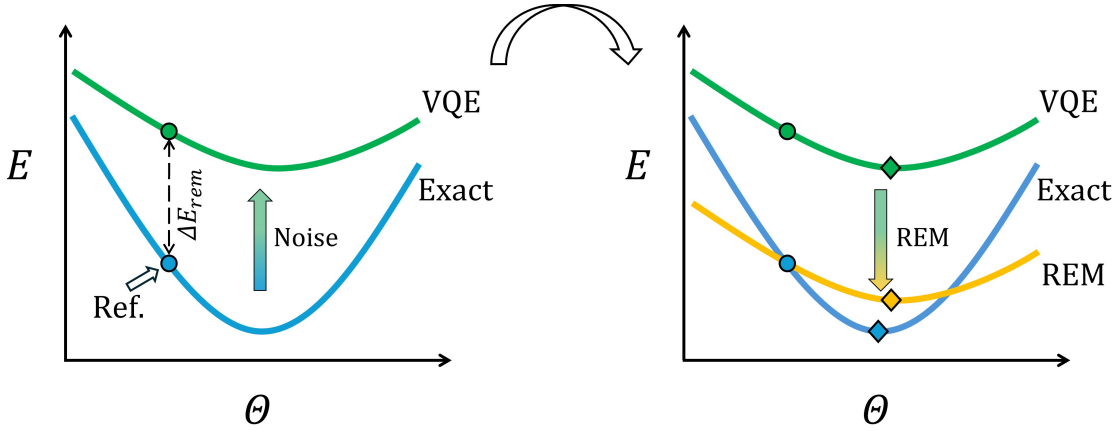


Figure 2.3: A strategy representation of the REM approach. Noise tends to flatten and elevate the energy landscape. Using a reference energy can partially capture the impact of noise and significantly mitigate the results of noisy VQE.

In REM, if the reference state is used as the initial state for the VQE experiment, no additional quantum cost is incurred. Ref. [99] employed the simplest HF reference state for error mitigation of the ground state energy of small molecules such as H_2 and HeH^+, resulting in a two orders of magnitude improvement in computational accuracy. This approach is chemically motivated, as the HF reference state is the starting point for all wavefunction theories and normally overlaps substantially with the ground state, making it a reliable choice. In contrast, random references made from Clifford groups are almost guaranteed to be ineffective.

2.4 Multireference-state error mitigation

2.4.1 Motivation

The development of multireference-state error mitigation (MR-REM) method aims to address the deficiencies of single-determinant descriptions in the original REM,

particularly evident in quantum systems with significant electron correlation effects. In regions involving bond stretching or in strongly correlated systems, the inadequacies of using a single determinant (such as HF method) to describe the electronic wavefunction become apparent, leading to an overestimation of electronic energy due to inadequate treatment of electron correlation. Fig. 2.4 demonstrates that employing a multireference state, which consists of a linear combination of the HF determinant and significant excited configurations, improves both the overlap with the true ground state and the accuracy of energy calculations compared to the HF theory. By preparing multireference states on a quantum computer, MR-REM is designed to rectify the inherent limitations of single-reference methods, thus providing a more accurate estimate of noise impact on the true ground state.

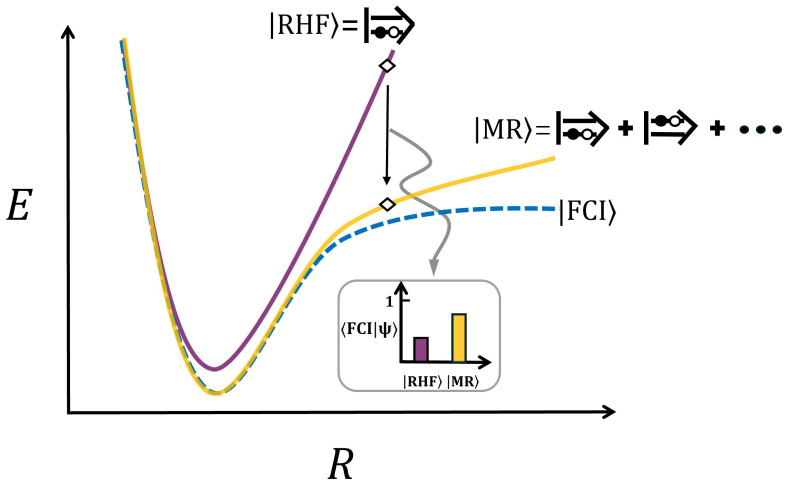


Figure 2.4: Multireference states approximate the ground state by considering a linear combination of multiple SDs.

2.4.2 The MR-REM procedure

The MR-REM method is summarized in Fig. 2.5. First, the configurations and coefficients of the truncated multireference wavefunction can be obtained from conventional quantum chemistry methods such as CISD, CCSD, and DMRG. Subsequently, a Givens gate-based multireference state preparation circuit is employed to prepare the multireference state on a quantum computer. Detailed information about Givens gates will be provided in the following section. The multireference state preparation circuit is designed to be gate-efficient, allowing it to be classically simulable and enabling us to determine the exact energy of the prepared multireference state. Following this, a parameterized ansätze circuit is attached to the multireference state preparation circuit. We sample and measure the noisy energy of the reference state using the complete circuit. The same circuits are then used to run the VQE algorithm and perform REM in the same steps.

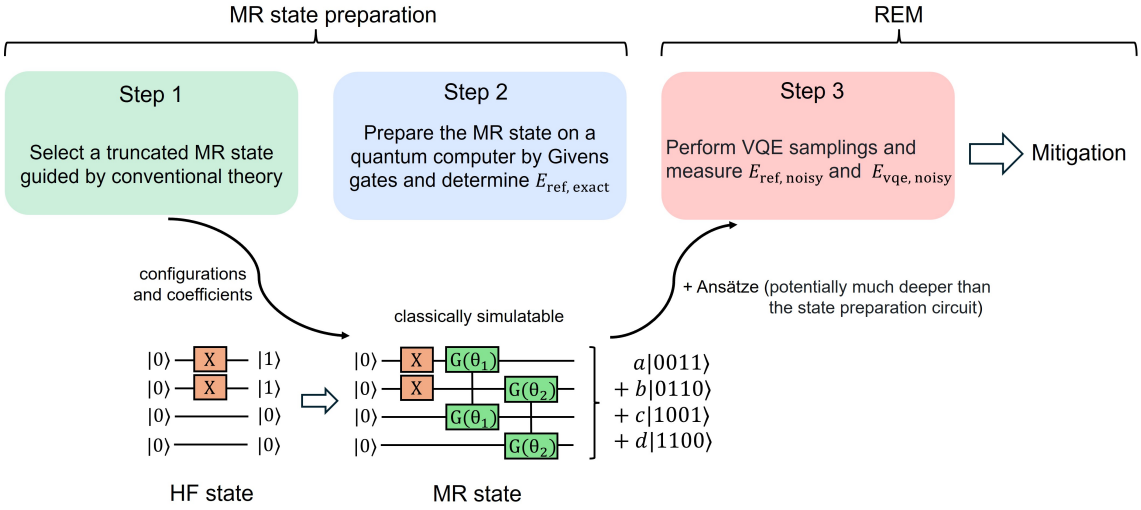


Figure 2.5: Overview of MR-REM: comprising multireference state preparation and reference state error migration. Extending the preparation of HF state on a quantum computer to prepare multireference states using Givens gate-based circuits.

2.4.3 Quantum lego: multireference state preparation

Symmetry in quantum chemistry provides constraints and guidance for state preparation, without requiring access to every point in Hilbert space. General circuits that respect symmetry can avoid preparing quantum states in incorrect subspaces. Ref. [101] demonstrates that controlled single-excitation gates are universal for particle-conserving unitaries. A candidate for single-excitation particle-conserving gates is the A gate (Eq. (2.40)) in the ASWAP ansätze. However, due to its inability to realize the identity operator and possessing weaker systematic improvability, we focus here on another type of single-excitation gate known as Givens rotations [101].

The Givens rotation can be expressed as

$$\begin{aligned} G(\theta) &= \exp\left[-i\frac{\theta}{2}(Y\otimes X - X\otimes Y)\right] \\ &= \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & \cos(\theta/2) & -\sin(\theta/2) & 0 \\ 0 & \sin(\theta/2) & \cos(\theta/2) & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}, \end{aligned} \tag{2.49}$$

thus it can span the single-excitation subspace as

$$\begin{aligned} G(\theta)\,|01\rangle &= \cos(\theta/2)\,|01\rangle + \sin(\theta/2)\,|10\rangle\,, \\ G(\theta)\,|10\rangle &= \cos(\theta/2)\,|10\rangle - \sin(\theta/2)\,|01\rangle\,. \end{aligned} \tag{2.50}$$

In addition, the following equalities hold

$$[G(\theta), N_\alpha] = 0, \quad [G(\theta), N_\beta] = 0, \quad \left[G(\theta), S^2\right] = 0. \tag{2.51}$$

Four-qubit double-excitation gates $G^{(2)}(\theta)$ can also be considered,

$$\begin{aligned} G^{(2)}(\theta)\,|0011\rangle &= \cos(\theta/2)\,|0011\rangle + \sin(\theta/2)\,|1100\rangle\,, \\ G^{(2)}(\theta)\,|1100\rangle &= \cos(\theta/2)\,|1100\rangle - \sin(\theta/2)\,|0011\rangle\,. \end{aligned} \tag{2.52}$$

The excitations induced by $G(\theta)$ and $G^{(2)}$ are illustrated in Fig. 2.6. The gate decompositions for $G(\theta)$ and $G^{(2)}(\theta)$ are detailed in Appendix A.1.

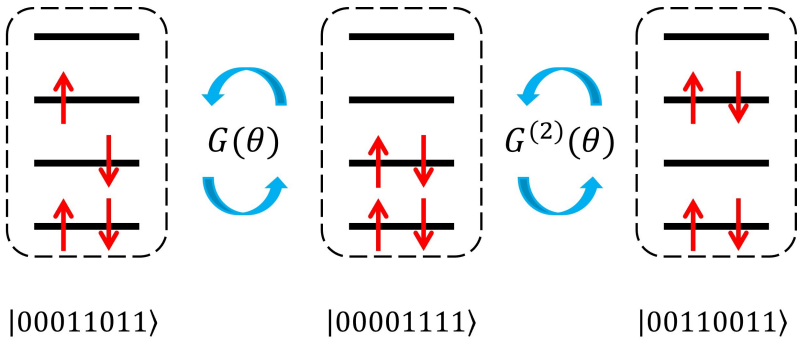


Figure 2.6: An example of $G(\theta)$ and $G^{(2)}(\theta)$ applied on a CAS(4e,4o) space. The qubits follow an interleaved spin ordering of electrons, i.e., $|\beta_3\alpha_3\beta_2\alpha_2\beta_1\alpha_1\beta_0\alpha_0\rangle$, where the rightmost qubit refers to the first qubit 0. $G(\theta)$ acts on qubits 4 and 2, achieving a single excitation of α_1 to α_2. $G^{(2)}(\theta)$ acts on qubits $4,5$ and $2,3$, achieving a double excitation of α_1, β_1 to α_2, β_2.

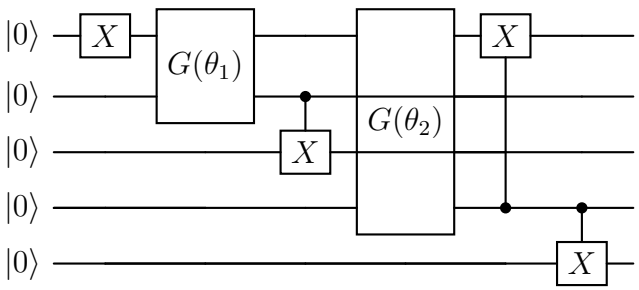
Additionally, the controlled single-excitation Givens gates can be written as

$$CG(\theta) = \begin{pmatrix} \mathbb{1}^{4\times4} & 0 \\ 0 & G(\theta) \end{pmatrix}, \tag{2.53}$$

where $\mathbb{1}^{4\times4}$ denotes a 4×4 identity matrix, and $CG(\theta)$ is a three-qubit (8×8) gate. Utilizing Givens gates and controlled Givens gates, arbitrary particle-preserving multireference states can be prepared. However, in other mappings such as the BK mapping or qubit tapering, the correspondence between particles and qubits is no longer one-to-one. In this case, multireference states can still be prepared using Givens gates, controlled Givens gates, and controlled-X gates. For instance, the state $a\,|00001\rangle + b\,|00110\rangle + c\,|11001\rangle$ can also represent a possible particle-preserving wavefunction under other mappings, which can be prepared by

$$CX_{3,4}\;CX_{3,0}\;G_{3,0}(\theta_2)\;CX_{1,2}\;G_{1,0}(\theta_1)\;X_0\,|00000\rangle$$

$$|||$$



, (2.54)

where the subscripts in $CX_{i,j}$ denote that i is the control qubit and j is the target qubit, and the rightmost qubit is denoted as the first qubit 0. The rotation parameters θ_1 and θ_2 can be determined by solving the equations

$$\begin{cases} a = \cos(\theta_1/2)\cos(\theta_2/2) \\ b = \sin(\theta_1/2) \\ c = \cos(\theta_1/2)\sin(\theta_2/2) \end{cases} \tag{2.55}$$

Apparently, there are multiple choices for constructing circuits to achieve the desired state, as well as multiple choices for obtaining parameters. A method to acquire the coefficients $(a, b, c, \cdots)$ and configurations of SDs is through the truncated wavefunction generated by conventional quantum chemistry methods [102]. The choice of multireferences similarly follows chemical motivation and leverages the advantages of hybrid algorithms. Upon obtaining the MR states that more accurately capture electronic correlation and exhibit increased overlap with the ground state, they can be leveraged to reduce the iteration count in the VQE algorithm and facilitate more efficient REM treatments.

3
Methods

In this chapter, we introduce the computational details and software packages used in this work, including the preparation of electronic Hamiltonians, initialization of multireference states, implementation of the VQE algorithm along with noisy simulations, and resource statistics of quantum circuits.

3.1 Computational details

The second quantized electronic Hamiltonian was obtained via the PySCF [103] library. For H_2, H_2O, and F_2, restricted Hartree-Fock (RHF) orbitals were computed using the double-ζ correlation-consistent basis set, cc-pVDZ, while for N_2, we employed the cc-pVTZ basis set. The `ActiveSpaceTransformer` in Qiskit [104] was used to extract the active orbital electronic Hamiltonian. The selection of active electrons and active orbitals was as follows: H_2 $(2e, 2o)$, H_2O $(4e, 4o)$, N_2 $(6e, 6o)$, and F_2 $(10e, 6o)$. All fermionic Hamiltonians were mapped to qubit form using the Jordan-Wigner transformation. Except for H_2, qubit tapering was applied to reduce the number of qubits for the other molecules: H_2O $(8 \to 5)$, N_2 $(12 \to 8)$, and F_2 $(12 \to 8)$. The exact diagonalization of qubit Hamiltonians was performed using the NumPy minimum eigensolver algorithm.

The selections of MR determinants are provided by classical theories, including CISD, CCSD, and DMRG. CISD and CCSD calculations were performed using the PySCF library, and DMRG calculations were carried out using the block2 code [105]. The `import_state` function in PennyLane [106] was used to convert an external wavefunction object into a state vector, thereby facilitating the extraction of SDs and their coefficients.

3.1.1 Noisy simulation

The implementation of all quantum circuits and the VQE algorithm was based on the `Estimator` module of Qiskit Aer 0.13.1. For all Estimator simulations, expectation values were computed by performing 1×10^7 shots of sampling. This extensive sampling was conducted to minimize statistical errors. In standard VQE, $\mathcal{O}(1/\epsilon^2) \sim 10^6$ shots are required to achieve an error rate of $\epsilon \sim 10^{-3}$ for small molecules. In addition, the `approximation` option of the Estimator was employed, leading to a notable enhancement of simulation efficiency by estimating the expecta-

tion value with sampling noise using a normal distribution approximation. However, it is important to acknowledge that this approach ignores readout errors. Therefore, in these experiments, the primary focus of our demonstrations is the migration simulation of quantum circuit gate errors.

The FakeSydneyV2 backend noise model was employed to provide noise simulations. IBM provides a noise model encompassing the depolarization and thermal relaxation errors of single- and two-qubit gates. These error parameters are derived from backend error parameters, including average gate errors of each basis gate, gate duration, single-qubit readout error probability, and T1 and T2 decoherence times for each qubit [107]. These parameters were built through access to real noise data for various IBM quantum processors.

The classical optimizer applied to VQE is the implicit filtering (ImFil) optimizer, implemented in the scikit-quant package [108]. ImFil is an optimization algorithm tailored for noisy problems with many local minima. It evaluates the objective function at multiple points within defined bounds and uses the best result to guide the next search. This approach allows it to handle noise effectively and avoid being trapped in local minima, making it robust and efficient for optimizing complex, noisy functions.

3.2 Resource statistics of quantum circuits

In the VQE experiments, we adopt the hardware-efficient R_Y ansatz to minimize circuit depth. The complete circuit structure is depicted in Fig. 3.1. The R_Y ansatz is placed in front of the MR preparation circuit because it satisfies $U(0)\,|0\rangle = |0\rangle$, while it is non-trivial to find parameters that allow it to act as an identity.

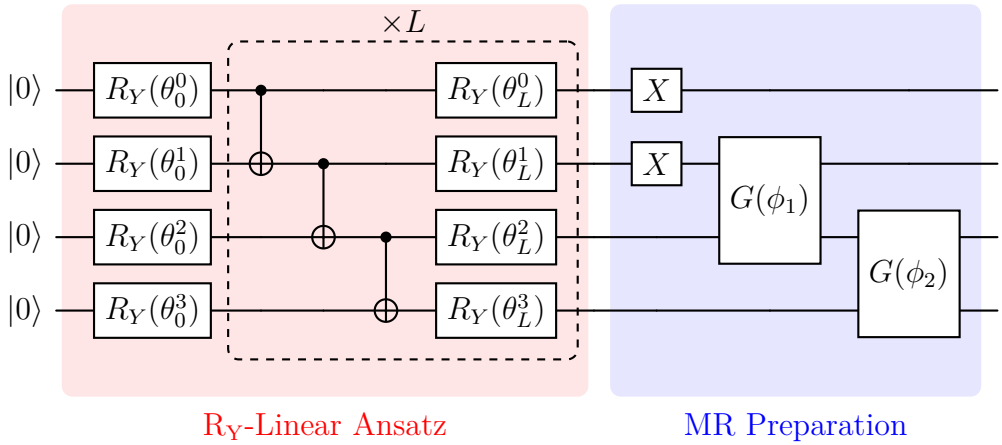


Figure 3.1: Schematic diagram of the complete circuit structure of MR-REM. All optimizable parameters $\boldsymbol{\theta}$ in the hardware-efficient R_Y ansatz are initialized to 0. The parameters $\boldsymbol{\phi}$ in the MR preparation circuit are obtained from a trial wavefunction through classical theory, which can be fixed or adjustable.

The information regarding the number of single-qubit gates and two-qubit gates in

the ansatz and MR preparation circuit is collected in Table 3.1. For a hardware-efficient R_Y ansatz, the number of optimizable parameters equals the number of single-qubit gates used. For H_2 and N_2, the application of $G^{(2)}(\theta)$ leads to relatively less gate-efficient MR preparation. In principle, a combination of $G(\theta)$ and $CG(\theta)$ can replace $G^{(2)}(\theta)$ to appropriately reduce the number of gates required.

Table 3.1: Statistics of Qubit Counts (N_{qb}), Layers Numbers (N_L), and Single (N_1) and Two-Qubit Gates (N_2) in the R_Y-Linear Ansatz and MR Preparation Circuits for Different Molecules. For N_1 in MR Circuit: The Numbers in Parentheses Represent the Gates Used in HF State Preparation.

molecule	N_{qb}	R_Y ansatz			MR circuit	
		N_L	N_1	N_2	N_1	N_2
H_2	4	1	8	3	16(2)	14
H_2O	5	5	30	20	9(1)	7
N_2	8	20	168	140	57(3)	42
F_2	8	5	48	35	10(6)	3

4

Results

4.1 Potential energy surface

4.1.1 $H_2(2e, 2o)$ and $F_2(10e, 6o)$

We start by investigating the performance improvement of MR-REM compared to single-reference REM for diatomic single-bond molecules H_2 and F_2. Such molecules can be well described by a linear combination of 2SDs in the bond stretching region. Therefore, we use quantum circuits to prepare a linear combination of 2 SDs as references for both molecules. The coefficients of the 2SDs are determined by the CISD method at the minimal basis set level. It is unnecessary for the reference state to be obtained at the same level of basis set as the target molecular Hamiltonian problem, as we only require a good approximation of the coefficients. For clarity, the VQE result with a HF reference is labeled as VQE-HF, and the corresponding REM result is denoted as REM-HF. Similarly, the VQE result using a linear combination of N SDs is labeled as VQE-NSDs, and its corresponding MR-REM result is denoted as REM-NSDs.

For the H_2 molecule, using only a one-layer R_Y ansatz results in insignificant gate noise, whereas the state preparation circuit using Givens rotations introduces large additional noise. Consequently, the energy of the VQE-2SDs is much higher than that of the VQE-HF, as shown in Fig. 4.1. However, all REM-2SDs tests display better results that reach computational accuracy. In contrast, the REM-HF fails to reach computational accuracy in the bond stretching region. Here, we prioritize computational accuracy over chemical accuracy, where the former refers to the error between the computational result and the exact value at the same level of theory, while the latter pertains to the accuracy required for realistic chemical predictions, as suggested in Ref. [99].

As illustrated in Fig. 4.1, it is evident for F_2 molecule that the REM-2SDs results improve by approximately two orders of magnitude compared to the noisy VQE results, and by one order of magnitude compared to the REM-HF results. For F_2, the additional noise introduced by the Givens circuit is tiny compared to the R_Y ansatz. Despite the fact that the VQE-2SDs circuit in noisy VQE introduces some additional noise compared to the VQE-HF circuit, it still yields a better optimized ground state result. In regions where the bond length R exceeds 1.5 Å, all REM-2SDs tests achieve computational accuracy. However, noisy VQE encounters difficulties in optimization in regions where R is smaller than 1.5 Å, likely due to being trapped in

local minima of the cost landscape. Nevertheless, a multireference state can alleviate this issue by offering a better starting point.

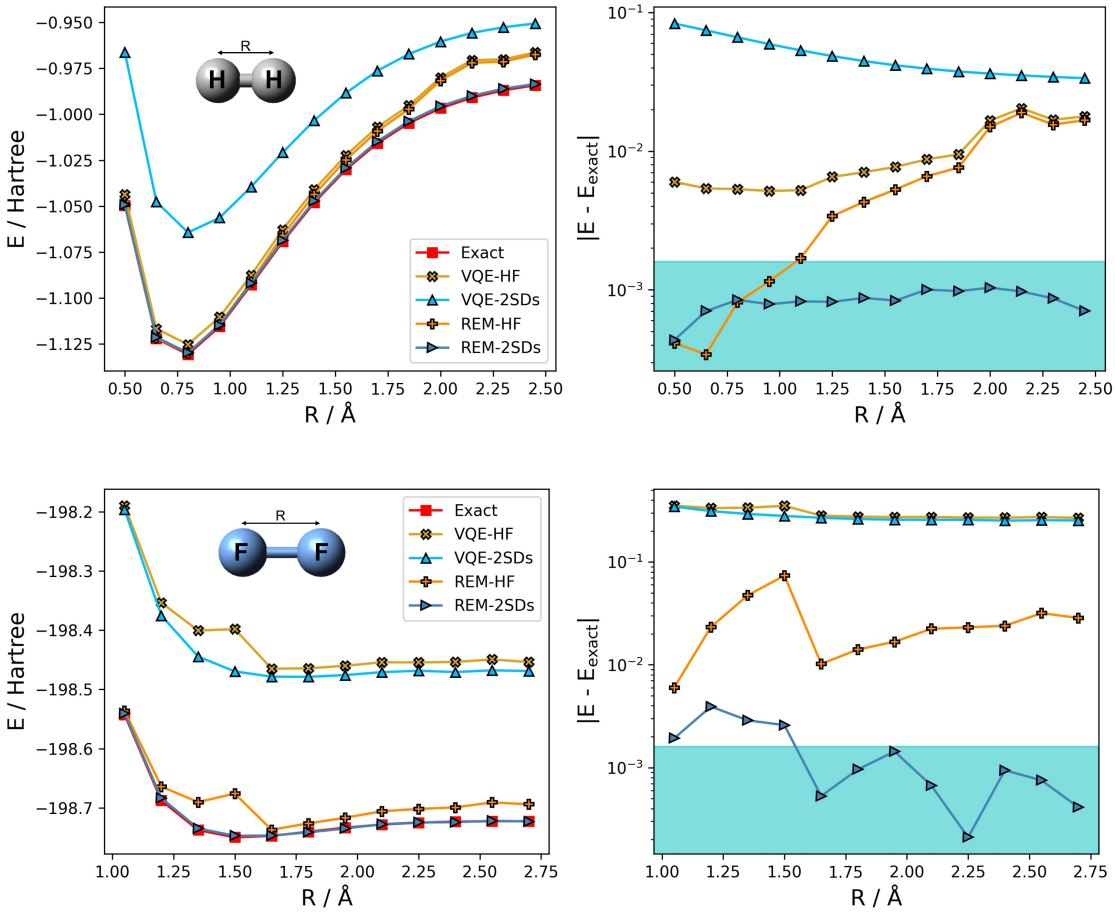


Figure 4.1: Comparisons of REM-2SDs and REM-HF for the potential energy surfaces (PES) of $H_2(2e, 2o)$ and $F_2(10e, 6o)$ at the cc-pVDZ level. The multireference 2SDs is obtained from truncated wavefunctions using the CISD method using a minimal STO-6G basis set. Left: PES for exact results, noisy VQE results with reference states of HF and 2SDs, and the corresponding REM results individually. Right: The absolute error of VQE and REM results with respect to the exact results. The cyan dashed lines represent the computational accuracy below 1.6×10^{-3} hartree (1 kcal/mol).

4.1.2 $H_2O(4e, 4o)$ and $N_2(6e, 6o)$

For the H_2O molecule, we prepare a linear combination of 3 SDs as a reference. Additionally, we introduce a penalty term that shifts the original Hamiltonian by $H' = H + 0.1S^2$ to prevent convergence to an undesired symmetry sector. Here, S^2 denotes the total spin angular momentum operator. Referring to Eq. 2.39, $S_i = S^2$, $s_i = 0$, and $\lambda_i = 0.1$. The only distinction is that the penalty term lacks the square operation. Since we aim to access the singlet ground state where $S^2 = 0$, any eigenvalue of S^2 would be non-negative. Hence, evaluating the square term is

unnecessary in this case. Moreover, avoiding the square term prevents the generation of additional Pauli terms, thus reducing measurement overhead. Similar to the case of F_2, the MR preparation for H_2O is gate-efficient (see Table 3.1), thus introducing only a small amount of additional noise. As a result, the VQE-3SDs energy is only slightly higher than that of VQE-HF, as shown in Fig. 4.2. In the bond stretching region, the results of REM-3SDs exhibit a twofold improvement over REM-HF, with errors exhibiting a converging trend.

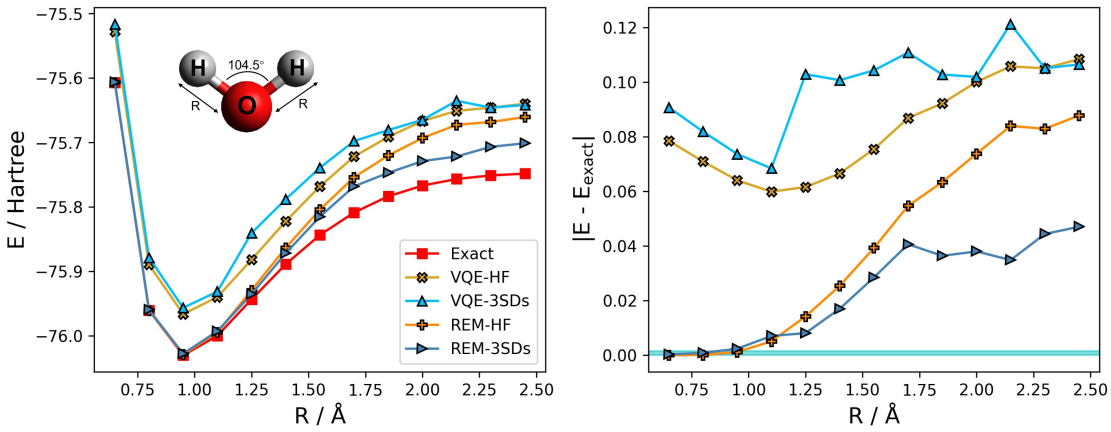


Figure 4.2: Comparison of REM-3SDs and REM-HF for the PES of $H_2O(4e, 4o)$ at the cc-pVDZ level. The PES scan varies both HO bond lengths R simultaneously, with a fixed $\angle HOH = 104.5°$. The multireference state of 3SDs is guided by CCSD/6-31G.

For the strongly correlated N_2 molecule, we introduce the multireference state of 3SDs guided by a DMRG estimate. Similar to the H_2O case, a penalty term of $H' = H + 0.5S^2$ is employed. As shown in Table 3.1, the circuits for VQE-3SDs use a relatively large number of additional gates, which introduces significant noise. Consequently, the VQE-HF curve outperforms the VQE-3SDs curve regarding both energy accuracy and the shape of the PES, as shown in Fig 4.3. However, the REM-3SDs generate a more convergent PES in the bond-stretching region, which is consistent with the asymptotic behaviour of the molecular potential. The results of REM-3SDs achieve significant improvements, with all tests showing absolute errors within approximately 0.1 Hartree. Additionally, the REM-3SDs curve enables the analysis of the equilibrium bond length region, which is indistinguishable in the VQE-3SDs curve.

In principle, increasing the number of determinants used for the multireference states can further improve the ground state approximation. For shallow circuits, even if additional noise is introduced, the improved mitigation results can be worthwhile. However, the extent of noise-induced disruption to the energy landscape is challenging to estimate. If excessive noise leads to noise-induced barren plateaus, the results become unreliable. Therefore, it is essential to consider the trade-off between the potential gains from increasing the number of reference determinants and the additional noise introduced.

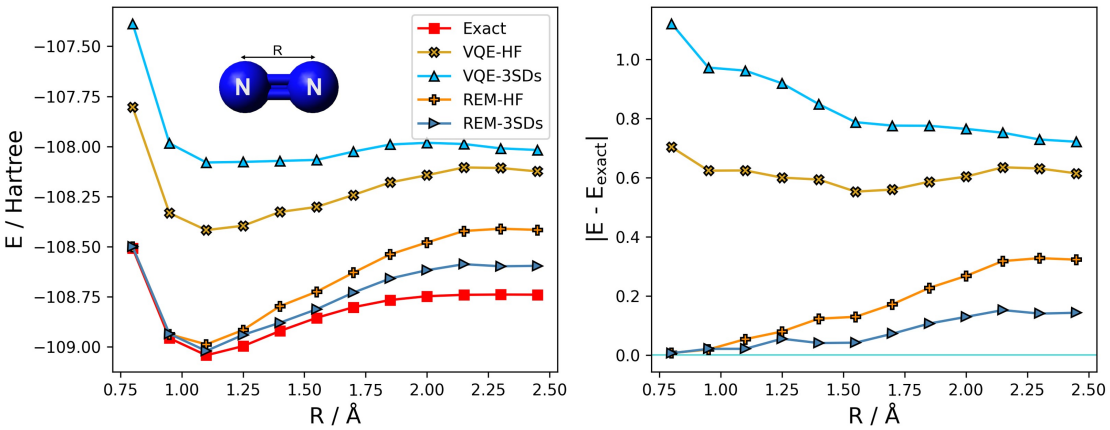


Figure 4.3: Comparison of REM-3SDs and REM-HF for the PES of $N_2(6e, 6o)$ at the cc-pVTZ level. The multireference state of 3SDs is guided by DMRG/cc-pVTZ.

4.2 Symmetry check

We check the spin symmetry by computing the expectation values of the observable S^2 for each noisy VQE ground state, $\langle\psi(\boldsymbol{\theta})|S^2|\psi(\boldsymbol{\theta})\rangle$. The results of H_2 and F_2 are shown in Figure 4.4. For these two molecules, a non-zero spin angular momentum implies the presence of a mixture of singlet and triplet states. This phenomenon arises from the combined effects of the hardware-efficient ansatz failing to respect symmetry and the presence of noise. The observed higher spin contamination in both REM-2SDs can be attributed to the increased noise in the circuits. Introducing constraints on symmetry, such as by adding penalty terms, does not mitigate this effect. This is because the addition of penalty terms often leads to more measurements, consequently amplifying the impact of noise. Fortunately, the VQE solution remains essentially within the singlet subspace, and this slight spin contamination does not pose a significant challenge to the qualitative analysis, meaning that the shape of PES will not be significantly disrupted.

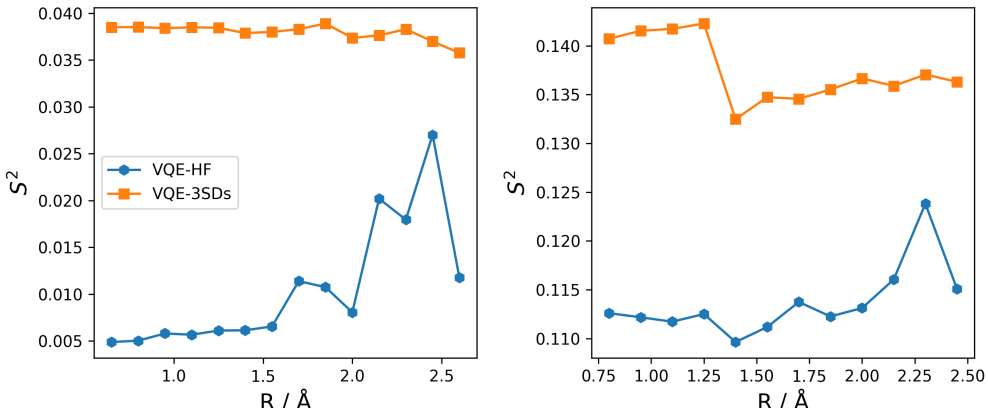


Figure 4.4: Spin angular momentum of noisy ground states of VQE-2SDs and VQE-HF for molecules. Left: $H_2(2e, 2o)$ and Right: $F_2(10e, 6o)$.

For the noisy VQE solutions of H_2O and N_2, noticeable symmetry breaking occurs, where the spin angular momentum undergoes sharp transitions, as shown in Fig. 4.5. This can result in non-differentiable effects on the PES. For these two molecules, the energies of the singlet and triplet states are close in the bond stretching region. Noise blurs the distinction between accessing global and local minima in the VQE cost landscape, causing VQE to erroneously tend towards accessing triplet states. Following the penalty scheme from the previous section, symmetry is noticeably restored. However, some noise-induced residuals remain challenging to avoid.

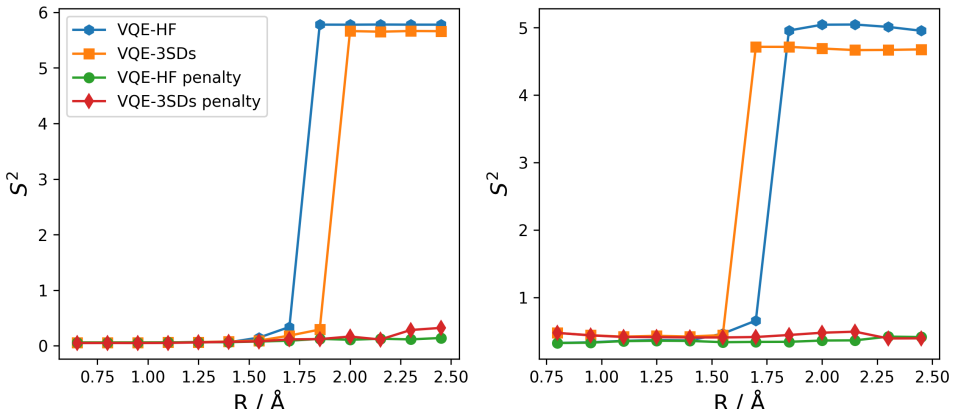


Figure 4.5: Angular momentum of noisy ground states of VQE-HF, VQE-3SDs, and their corresponding solutions with penalty terms for molecules. Left: $H_2O(4e, 4o)$ and Right: $N_2(6e, 6o)$. The Hamiltonian with penalty terms is denoted as $H' = H + \lambda S^2$, where $\lambda = 0.1$ for H_2O and $\lambda = 0.5$ for N_2. The λS^2 is introduced to enforce proper constraints on the quantum states by penalizing deviations from the desired values of angular momentum.

5
Conclusion

In this thesis, we review the theoretical background on solving the electronic structure problem on quantum computers and quantum error mitigation methods. We exhibit the idea and method of multireference-state error mitigation. The key task is the preparation of high-quality multireference states on a quantum computer, for which Givens rotations provide an efficient approach. The improvement achieved by MR-REM has been demonstrated through digital quantum simulations, with advantages including two main aspects. Firstly, it provides a better starting point for VQE, reducing the risk of becoming trapped in a local minimum and the number of VQE optimization iterations. Secondly, it achieves a better overlap with the target ground state, allowing for more effective noise mitigation and reducing the impact of noise on parameter shifts.

MR-REM shows promise as a potential method for dealing with strongly correlated molecules. However, we are still facing a dilemma in the era of NISQ. As seen in our example of N_2, the depth of the circuit has already caused obvious disruptions in the PES. For more correlated molecules, QEC or mitigation methods like PEC may have a positive effect on computational results but come with heavy costs in terms of quantum resources and time consumption. MR-REM offers a cost-efficient solution, but just as the no free lunch theorem implies, it is still limited for treating larger systems since it does not alleviate some inherent challenges in VQE, such as $\mathcal{O}(N^4)$ measurements, NP-hard optimization of variational parameters [109], and noise-induced barren plateaus. In practical application, the combination of MR-REM and other methods is anticipated to yield remarkable outcomes, such as the incorporation of measurement protocols like commuting groups [110] and classical shadows [111]. In the examples provided, a linear combination of at most three SDs prepared by Givens circuits are used as a reference. However, attempting to prepare a larger number of SDs would not be hardware-efficient. Hence, there is a demand for a more circuit-compact approach to represent quantum states on a quantum computer. Most multireference preparation methods based on fragmented QPE are not suitable for near-term devices. One potentially effective representation is the quantum computing version of neural network quantum states (NQS). NQS is considered to have higher expressive power than tensor networks [112], and their quantum computing versions potentially provide more efficient multireference approximations.

In general, although NISQ quantum algorithms combined with quantum error mitigation have made significant progress in practice, there is still a considerable gap

to surpass classical algorithms. With ongoing advancements in hardware and algorithmic development, we are moving towards the ability to compute larger active spaces with high precision.

Bibliography

[1] Paul Adrien Maurice Dirac. Quantum mechanics of many-electron systems. *Proceedings of the Royal Society of London. Series A, Containing Papers of a Mathematical and Physical Character*, 123(792):714–733, 1929.

[2] Richard P Feynman. Simulating physics with computers. In *Feynman and computation*, pages 133–153. CRC Press, 2018.

[3] Peter W Shor. Algorithms for quantum computation: discrete logarithms and factoring. In *Proceedings 35th annual symposium on foundations of computer science*, pages 124–134. Ieee, 1994.

[4] Christopher Portmann and Renato Renner. Security in quantum cryptography. *Reviews of Modern Physics*, 94(2):025008, 2022.

[5] Chao-Yang Lu, Yuan Cao, Cheng-Zhi Peng, and Jian-Wei Pan. Micius quantum experiments in space. *Reviews of Modern Physics*, 94(3):035001, 2022.

[6] Christian L Degen, Friedemann Reinhard, and Paola Cappellaro. Quantum sensing. *Reviews of modern physics*, 89(3):035002, 2017.

[7] Lena Funcke, Tobias Hartung, Karl Jansen, and Stefan Kühn. Review on quantum computing for lattice field theory. *arXiv preprint arXiv:2302.00467*, 2023.

[8] Junyu Liu and Yuan Xin. Quantum simulation of quantum field theories as quantum chemistry. *Journal of High Energy Physics*, 2020(12):1–48, 2020.

[9] Sam McArdle, Suguru Endo, Alán Aspuru-Guzik, Simon C Benjamin, and Xiao Yuan. Quantum computational chemistry. *Reviews of Modern Physics*, 92(1):015003, 2020.

[10] Iulia M Georgescu, Sahel Ashhab, and Franco Nori. Quantum simulation. *Reviews of Modern Physics*, 86(1):153, 2014.

[11] Yudong Cao, Jonathan Romero, Jonathan P Olson, Matthias Degroote, Peter D Johnson, Mária Kieferová, Ian D Kivlichan, Tim Menke, Borja Peropadre, Nicolas PD Sawaya, et al. Quantum chemistry in the age of quantum computing. *Chemical reviews*, 119(19):10856–10915, 2019.

[12] Non-abelian braiding of graph vertices in a superconducting processor. *Nature*, 618(7964):264–269, 2023.

[13] Antonis Kyprianidis, Francisco Machado, William Morong, Patrick Becker, Kate S Collins, Dominic V Else, Lei Feng, Paul W Hess, Chetan Nayak, Guido Pagano, et al. Observation of a prethermal discrete time crystal. *Science*, 372(6547):1192–1196, 2021.

[14] Damiano Aliverti-Piuri, Kaustav Chatterjee, Lexin Ding, Ke Liao, Julia Liebert, and Christian Schilling. What can quantum information theory offer to quantum chemistry? *Faraday Discussions*, 2024.

[15] Bei Zeng, Xie Chen, Duan-Lu Zhou, Xiao-Gang Wen, et al. *Quantum information meets quantum matter*. Springer, 2019.

[16] Ahmed Almheiri, Xi Dong, and Daniel Harlow. Bulk locality and quantum error correction in ads/cft. *Journal of High Energy Physics*, 2015(4):1–34, 2015.

[17] Fernando Pastawski, Beni Yoshida, Daniel Harlow, and John Preskill. Holographic quantum error-correcting codes: Toy models for the bulk/boundary correspondence. *Journal of High Energy Physics*, 2015(6):1–55, 2015.

[18] Robert J Harris, Elliot Coupe, Nathan A McMahon, Gavin K Brennen, and Thomas M Stace. Decoding holographic codes with an integer optimization decoder. *Physical Review A*, 102(6):062417, 2020.

[19] Sam Cree, Kfir Dolev, Vladimir Calvera, and Dominic J Williamson. Fault-tolerant logical gates in holographic stabilizer codes are severely restricted. *PRX Quantum*, 2(3):030337, 2021.

[20] Mark Horowitz and Emily Grumbling. Quantum computing: progress and prospects. 2019.

[21] Nathalie P De Leon, Kohei M Itoh, Dohun Kim, Karan K Mehta, Tracy E Northup, Hanhee Paik, BS Palmer, Nitin Samarth, Sorawis Sangtawesin, and David W Steuerman. Materials challenges and opportunities for quantum computing hardware. *Science*, 372(6539):eabb2823, 2021.

[22] Hans J Briegel, David E Browne, Wolfgang Dür, Robert Raussendorf, and Maarten Van den Nest. Measurement-based quantum computation. *Nature Physics*, 5(1):19–26, 2009.

[23] Tameem Albash and Daniel A Lidar. Adiabatic quantum computation. *Reviews of Modern Physics*, 90(1):015002, 2018.

[24] Andrew M Childs. Universal computation by quantum walk. *Physical review letters*, 102(18):180501, 2009.

[25] A Yu Kitaev. Fault-tolerant quantum computation by anyons. *Annals of physics*, 303(1):2–30, 2003.

[26] Sebastian Deffner and Steve Campbell. *Quantum Thermodynamics: An introduction to the thermodynamics of quantum information*. Morgan & Claypool Publishers, 2019.

[27] John Goold, Marcus Huber, Arnau Riera, Lídia Del Rio, and Paul Skrzypczyk. The role of quantum information in thermodynamics—a topical review. *Journal of Physics A: Mathematical and Theoretical*, 49(14):143001, 2016.

[28] Sergey B Bravyi and A Yu Kitaev. Quantum codes on a lattice with boundary. *arXiv preprint quant-ph/9811052*, 1998.

[29] Dolev Bluvstein, Simon J Evered, Alexandra A Geim, Sophie H Li, Hengyun Zhou, Tom Manovitz, Sepehr Ebadi, Madelyn Cain, Marcin Kalinowski, Dominik Hangleiter, et al. Logical quantum processor based on reconfigurable atom arrays. *Nature*, 626(7997):58–65, 2024.

[30] MP da Silva, C Ryan-Anderson, JM Bello-Rivas, A Chernoguzov, JM Dreiling, C Foltz, JP Gaebler, TM Gatterman, D Hayes, N Hewitt, et al. Demonstration of logical qubits and repeated error correction with better-than-physical error rates. *arXiv preprint arXiv:2404.02280*, 2024.

[31] Craig Gidney and Martin Ekerå. How to factor 2048 bit rsa integers in 8 hours using 20 million noisy qubits. *Quantum*, 5:433, 2021.

[32] Markus Reiher, Nathan Wiebe, Krysta M Svore, Dave Wecker, and Matthias Troyer. Elucidating reaction mechanisms on quantum computers. *Proceedings of the national academy of sciences*, 114(29):7555–7560, 2017.

[33] Zhendong Li, Junhao Li, Nikesh S Dattani, CJ Umrigar, and Garnet Kin Chan. The electronic complexity of the ground-state of the femo cofactor of nitrogenase as relevant to quantum simulations. *The Journal of chemical physics*, 150(2), 2019.

[34] Edward Farhi, Jeffrey Goldstone, and Sam Gutmann. A quantum approximate optimization algorithm. *arXiv preprint arXiv:1411.4028*, 2014.

[35] Alberto Peruzzo, Jarrod McClean, Peter Shadbolt, Man-Hong Yung, Xiao-Qi Zhou, Peter J Love, Alán Aspuru-Guzik, and Jeremy L O'brien. A variational eigenvalue solver on a photonic quantum processor. *Nature communications*, 5(1):4213, 2014.

[36] Kishor Bharti, Alba Cervera-Lierta, Thi Ha Kyaw, Tobias Haug, Sumner Alperin-Lea, Abhinav Anand, Matthias Degroote, Hermanni Heimonen, Jakob S Kottmann, Tim Menke, et al. Noisy intermediate-scale quantum algorithms. *Reviews of Modern Physics*, 94(1):015004, 2022.

[37] John Preskill. Quantum computing 40 years later. In *Feynman Lectures on Computation*, pages 193–244. CRC Press, 2023.

[38] Alán Aspuru-Guzik, Anthony D Dutoi, Peter J Love, and Martin Head-Gordon. Simulated quantum computation of molecular energies. *Science*, 309(5741):1704–1707, 2005.

[39] Laura Clinton, Johannes Bausch, Joel Klassen, and Toby Cubitt. Phase estimation of local hamiltonians on nisq hardware. *New Journal of Physics*, 25(3):033027, 2023.

[40] Hamed Mohammadbagherpoor, Young-Hyun Oh, Patrick Dreher, Anand Singh, Xianqing Yu, and Andy J Rindos. An improved implementation approach for quantum phase estimation on quantum computers. In *2019 IEEE International Conference on Rebooting Computing (ICRC)*, pages 1–9. IEEE, 2019.

[41] Seunghoon Lee, Joonho Lee, Huanchen Zhai, Yu Tong, Alexander M Dalzell, Ashutosh Kumar, Phillip Helms, Johnnie Gray, Zhi-Hao Cui, Wenyuan Liu, et al. Evaluating the evidence for exponential quantum advantage in ground-state quantum chemistry. *Nature communications*, 14(1):1952, 2023.

[42] Oscar Higgott, Daochen Wang, and Stephen Brierley. Variational quantum computation of excited states. *Quantum*, 3:156, 2019.

[43] Cristian L Cortes and Stephen K Gray. Quantum krylov subspace algorithms for ground-and excited-state energy estimation. *Physical Review A*, 105(2):022417, 2022.

[44] Pauline J Ollitrault, Abhinav Kandala, Chun-Fu Chen, Panagiotis Kl Barkoutsos, Antonio Mezzacapo, Marco Pistoia, Sarah Sheldon, Stefan Woerner, Jay M Gambetta, and Ivano Tavernelli. Quantum equation of motion for computing molecular excitation energies on a noisy quantum processor. *Physical Review Research*, 2(4):043140, 2020.

[45] Mario Motta, Chong Sun, Adrian TK Tan, Matthew J O'Rourke, Erika Ye, Austin J Minnich, Fernando GSL Brandao, and Garnet Kin-Lic Chan. Determining eigenstates and thermal states on a quantum computer using quantum imaginary time evolution. *Nature Physics*, 16(2):205–210, 2020.

[46] Hirsh Kamakari, Shi-Ning Sun, Mario Motta, and Austin J Minnich. Digital quantum simulation of open quantum systems using quantum imaginary–time evolution. *PRX Quantum*, 3(1):010320, 2022.

[47] Guglielmo Mazzola, Pauline J Ollitrault, Panagiotis Kl Barkoutsos, and Ivano Tavernelli. Nonunitary operations for ground-state calculations in near-term quantum computers. *Physical review letters*, 123(13):130501, 2019.

[48] Werner Dobrautz, Igor O Sokolov, Ke Liao, Pablo López Ríos, Martin Rahm, Ali Alavi, and Ivano Tavernelli. Toward real chemical accuracy on current quantum hardware through the transcorrelated method. *Journal of Chemical Theory and Computation*, 2024.

[49] Erika Magnusson, Aaron Fitzpatrick, Stefan Knecht, Martin Rahm, and Werner Dobrautz. Towards efficient quantum computing for quantum chemistry: Reducing circuit complexity with transcorrelated and adaptive ansatz techniques. *arXiv preprint arXiv:2402.16659*, 2024.

[50] Shi-Xin Zhang, Zhou-Quan Wan, Chee-Kong Lee, Chang-Yu Hsieh, Shengyu Zhang, and Hong Yao. Variational quantum-neural hybrid eigensolver. *Physical Review Letters*, 128(12):120502, 2022.

[51] Bruno Senjean, Saad Yalouz, and Matthieu Saubanère. Toward density functional theory on quantum computers? *SciPost Physics*, 14(3):055, 2023.

[52] Karl Michael Ziems, Erik Rosendahl Kjellgren, Peter Reinholdt, Phillip WK Jensen, Stephan PA Sauer, Jacob Kongsted, and Sonia Coriani. Which options exist for nisq-friendly linear response formulations? *Journal of Chemical Theory and Computation*, 2024.

[53] Manas Sajjan, Junxu Li, Raja Selvarajan, Shree Hari Sureshbabu, Sumit Suresh Kale, Rishabh Gupta, Vinit Singh, and Sabre Kais. Quantum machine learning for chemistry and physics. *Chemical Society Reviews*, 51(15):6475–6573, 2022.

[54] Max Born and Robert Oppenheimer. On the quantum theory of molecules. In *Quantum Chemistry: Classic Scientific Papers*, pages 1–24. World Scientific, 2000.

[55] Graham A Worth and Lorenz S Cederbaum. Beyond born-oppenheimer: Molecular dynamics through a conical intersection. *Annu. Rev. Phys. Chem.*, 55:127–158, 2004.

[56] Attila Szabo and Neil S Ostlund. *Modern quantum chemistry: introduction to advanced electronic structure theory*. Courier Corporation, 2012.

[57] Jack Simons. Why is quantum chemistry so complicated? *Journal of the American Chemical Society*, 145(8):4343–4354, 2023.

[58] Balazs Nagy and Frank Jensen. Basis sets in quantum chemistry. *Reviews in Computational Chemistry*, 30:93–149, 2017.

[59] Alexis Philip Ralli. *Practical quantum chemistry on near term quantum computers*. PhD thesis, UCL (University College London), 2023.

[60] Konstantinos D Vogiatzis, Dongxia Ma, Jeppe Olsen, Laura Gagliardi, and Wibe A De Jong. Pushing configuration-interaction to the limit: Towards massively parallel mcscf calculations. *The Journal of chemical physics*, 147(18), 2017.

[61] Steven R White. Density matrix formulation for quantum renormalization groups. *Physical review letters*, 69(19):2863, 1992.

[62] George H Booth, Alex JW Thom, and Ali Alavi. Fermion monte carlo without fixed nodes: A game of life, death, and annihilation in slater determinant space. *The Journal of chemical physics*, 131(5), 2009.

[63] Emmanuel Giner, Anthony Scemama, and Michel Caffarel. Using perturbatively selected configuration interaction in quantum monte carlo calculations. *Canadian Journal of Chemistry*, 91(9):879–885, 2013.

[64] Gerald Knizia and Garnet Kin-Lic Chan. Density matrix embedding: A simple alternative to dynamical mean-field theory. *Physical review letters*, 109(18):186404, 2012.

[65] Anton Frisk Kockum, Ariadna Soro, Laura García-Álvarez, Pontus Vikstål, Tom Douce, Göran Johansson, and Giulia Ferrini. Lecture notes on quantum computing. *arXiv preprint arXiv:2311.08445*, 2023.

[66] Frank Verstraete and J Ignacio Cirac. Mapping local hamiltonians of fermions to local hamiltonians of spins. *Journal of Statistical Mechanics: Theory and Experiment*, 2005(09):P09012, 2005.

[67] Mitchell Chiew and Sergii Strelchuk. Discovering optimal fermion-qubit mappings through algorithmic enumeration. *Quantum*, 7:1145, 2023.

[68] Jacob T Seeley, Martin J Richard, and Peter J Love. The bravyi-kitaev transformation for quantum computation of electronic structure. *The Journal of chemical physics*, 137(22), 2012.

[69] Jules Tilly, Hongxiang Chen, Shuxiang Cao, Dario Picozzi, Kanav Setia, Ying Li, Edward Grant, Leonard Wossnig, Ivan Rungger, George H Booth, et al. The variational quantum eigensolver: a review of methods and best practices. *Physics Reports*, 986:1–128, 2022.

[70] Nicolas PD Sawaya, Mikhail Smelyanskiy, Jarrod R McClean, and Alán Aspuru-Guzik. Error sensitivity to environmental noise in quantum circuits for chemical state preparation. *Journal of chemical theory and computation*, 12(7):3097–3108, 2016.

[71] Sergey Bravyi, Jay M Gambetta, Antonio Mezzacapo, and Kristan Temme. Tapering off qubits to simulate fermionic hamiltonians. *arXiv preprint arXiv:1701.08213*, 2017.

[72] Utkarsh Azad and Soran Jahangiri. Qubit tapering. `https://pennylane.ai/qml/demos/tutorial_qubit_tapering/`, 05 2022. Date Accessed: 2024-05-04.

[73] Joonho Lee, William J Huggins, Martin Head-Gordon, and K Birgitta Whaley. Generalized unitary coupled cluster wave functions for quantum computation. *Journal of chemical theory and computation*, 15(1):311–324, 2018.

[74] Abhinav Anand, Philipp Schleich, Sumner Alperin-Lea, Phillip WK Jensen, Sukin Sim, Manuel Díaz-Tinoco, Jakob S Kottmann, Matthias Degroote, Artur F Izmaylov, and Alán Aspuru-Guzik. A quantum computing view on

unitary coupled cluster theory. *Chemical Society Reviews*, 51(5):1659–1684, 2022.

[75] Abhinav Kandala, Antonio Mezzacapo, Kristan Temme, Maika Takita, Markus Brink, Jerry M Chow, and Jay M Gambetta. Hardware-efficient variational quantum eigensolver for small molecules and quantum magnets. *nature*, 549(7671):242–246, 2017.

[76] Jarrod R McClean, Sergio Boixo, Vadim N Smelyanskiy, Ryan Babbush, and Hartmut Neven. Barren plateaus in quantum neural network training landscapes. *Nature communications*, 9(1):4812, 2018.

[77] Zoë Holmes, Kunal Sharma, Marco Cerezo, and Patrick J Coles. Connecting ansatz expressibility to gradient magnitudes and barren plateaus. *PRX Quantum*, 3(1):010313, 2022.

[78] Lorenzo Leone, Salvatore FE Oliviero, Lukasz Cincio, and M Cerezo. On the practical usefulness of the hardware efficient ansatz. *arXiv preprint arXiv:2211.01477*, 2022.

[79] Xin Wang, Bo Qi, Yabo Wang, and Daoyi Dong. Entanglement-variational hardware-efficient ansatz for eigensolvers. *Physical Review Applied*, 21(3):034059, 2024.

[80] Chufan Lyu, Xusheng Xu, Man-Hong Yung, and Abolfazl Bayat. Symmetry enhanced variational quantum spin eigensolver. *Quantum*, 7:899, 2023.

[81] Ilya G Ryabinkin, Scott N Genin, and Artur F Izmaylov. Constrained variational quantum eigensolver: Quantum computer search engine in the fock space. *Journal of chemical theory and computation*, 15(1):249–255, 2018.

[82] Wusheng Zhu and Herschel Rabitz. A rapid monotonically convergent iteration algorithm for quantum optimal control over the expectation value of a positive definite operator. *The Journal of Chemical Physics*, 109(2):385–391, 1998.

[83] Bryan T Gard, Linghua Zhu, George S Barron, Nicholas J Mayhall, Sophia E Economou, and Edwin Barnes. Efficient symmetry-preserving state preparation circuits for the variational quantum eigensolver algorithm. *npj Quantum Information*, 6(1):10, 2020.

[84] Gian-Luca R Anselmetti, David Wierichs, Christian Gogolin, and Robert M Parrish. Local, expressive, quantum-number-preserving vqe ansätze for fermionic systems. *New Journal of Physics*, 23(11):113010, 2021.

[85] Hugh GA Burton. Accurate and gate-efficient quantum ansatze for electronic states without adaptive optimization. *arXiv preprint arXiv:2312.09761*, 2023.

[86] Enrico Fontana, M Cerezo, Andrew Arrasmith, Ivan Rungger, and Patrick J Coles. Non-trivial symmetries in quantum landscapes and their resilience to quantum noise. *Quantum*, 6:804, 2022.

[87] Samson Wang, Piotr Czarnik, Andrew Arrasmith, Marco Cerezo, Lukasz Cincio, and Patrick J Coles. Can error mitigation improve trainability of noisy variational quantum algorithms? *Quantum*, 8:1287, 2024.

[88] Jeff S Lundeen, Alvaro Feito, Hendrik Coldenstrodt-Ronge, Kenny L Pregnell, Ch Silberhorn, Timothy C Ralph, Jens Eisert, Martin B Plenio, and Ian A Walmsley. Tomography of quantum detectors. *Nature Physics*, 5(1):27–30, 2009.

[89] Bibek Pokharel, Siddarth Srinivasan, Gregory Quiroz, and Byron Boots. Scalable measurement error mitigation via iterative bayesian unfolding. *Physical Review Research*, 6(1):013187, 2024.

[90] Paul D Nation, Hwajung Kang, Neereja Sundaresan, and Jay M Gambetta. Scalable mitigation of measurement errors on quantum computers. *PRX Quantum*, 2(4):040326, 2021.

[91] Tudor Giurgica-Tiron, Yousef Hindy, Ryan LaRose, Andrea Mari, and William J Zeng. Digital zero noise extrapolation for quantum error mitigation. In *2020 IEEE International Conference on Quantum Computing and Engineering (QCE)*, pages 306–316. IEEE, 2020.

[92] Kristan Temme, Sergey Bravyi, and Jay M Gambetta. Error mitigation for short-depth quantum circuits. *Physical review letters*, 119(18):180509, 2017.

[93] Ewout Van Den Berg, Zlatko K Minev, Abhinav Kandala, and Kristan Temme. Probabilistic error cancellation with sparse pauli–lindblad models on noisy quantum processors. *Nature Physics*, 19(8):1116–1121, 2023.

[94] Senrui Chen, Yunchao Liu, Matthew Otten, Alireza Seif, Bill Fefferman, and Liang Jiang. The learnability of pauli noise. *Nature Communications*, 14(1):52, 2023.

[95] Emanuel Knill, Dietrich Leibfried, Rolf Reichle, Joe Britton, R Brad Blakestad, John D Jost, Chris Langer, Roee Ozeri, Signe Seidelin, and David J Wineland. Randomized benchmarking of quantum gates. *Physical Review A*, 77(1):012307, 2008.

[96] Heinz-Peter Breuer and Francesco Petruccione. *The theory of open quantum systems*. OUP Oxford, 2002.

[97] Zhenyu Cai, Ryan Babbush, Simon C Benjamin, Suguru Endo, William J Huggins, Ying Li, Jarrod R McClean, and Thomas E O'Brien. Quantum error mitigation. *Reviews of Modern Physics*, 95(4):045005, 2023.

[98] Xavi Bonet-Monroig, Ramiro Sagastizabal, M Singh, and TE O'Brien. Low-cost error mitigation by symmetry verification. *Physical Review A*, 98(6):062339, 2018.

[99] Phalgun Lolur, Mårten Skogh, Werner Dobrautz, Christopher Warren, Janka Biznárová, Amr Osman, Giovanna Tancredi, Goran Wendin, Jonas Bylander, and Martin Rahm. Reference-state error mitigation: A strategy for high accuracy quantum computation of chemistry. *Journal of Chemical Theory and Computation*, 19(3):783–789, 2023.

[100] Daniel Gottesman. The heisenberg representation of quantum computers. *arXiv preprint quant-ph/9807006*, 1998.

[101] Juan Miguel Arrazola, Olivia Di Matteo, Nicolás Quesada, Soran Jahangiri, Alain Delgado, and Nathan Killoran. Universal quantum circuits for quantum chemistry. *Quantum*, 6:742, 2022.

[102] Stepan Fomichev, Kasra Hejazi, Modjtaba Shokrian Zini, Matthew Kiser, Joana Fraxanet Morales, Pablo Antonio Moreno Casares, Alain Delgado, Joonsuk Huh, Arne-Christian Voigt, Jonathan E Mueller, et al. Initial state preparation for quantum chemistry on quantum computers. *arXiv preprint arXiv:2310.18410*, 2023.

[103] Qiming Sun, Xing Zhang, Samragni Banerjee, Peng Bao, Marc Barbry, Nick S Blunt, Nikolay A Bogdanov, George H Booth, Jia Chen, Zhi-Hao Cui, et al. Recent developments in the pyscf program package. *The Journal of chemical physics*, 153(2), 2020.

[104] Qiskit Community. Qiskit: An open-source framework for quantum computing, March 2017.

[105] Huanchen Zhai, Henrik R Larsson, Seunghoon Lee, Zhi-Hao Cui, Tianyu Zhu, Chong Sun, Linqing Peng, Ruojing Peng, Ke Liao, Johannes Tölle, et al. Block2: A comprehensive open source framework to develop and apply state-of-the-art dmrg algorithms in electronic structure and beyond. *The Journal of Chemical Physics*, 159(23), 2023.

[106] Ville Bergholm, Josh Izaac, Maria Schuld, Christian Gogolin, Shahnawaz Ahmed, Vishnu Ajith, M Sohaib Alam, Guillermo Alonso-Linaje, B Akash-Narayanan, Ali Asadi, et al. Pennylane: Automatic differentiation of hybrid quantum-classical computations. *arXiv preprint arXiv:1811.04968*, 2018.

[107] Christopher J Wood. Special session: Noise characterization and error mitigation in near-term quantum computers. In *2020 IEEE 38th International Conference on Computer Design (ICCD)*, pages 13–16. IEEE, 2020.

[108] Wim Lavrijsen, Ana Tudor, Juliane Müller, Costin Iancu, and Wibe De Jong. Classical optimizers for noisy intermediate-scale quantum devices. In *2020 IEEE international conference on quantum computing and engineering (QCE)*, pages 267–277. IEEE, 2020.

[109] Lennart Bittel and Martin Kliesch. Training variational quantum algorithms is np-hard. *Physical review letters*, 127(12):120502, 2021.

[110] Pranav Gokhale, Olivia Angiuli, Yongshan Ding, Kaiwen Gui, Teague Tomesh, Martin Suchara, Margaret Martonosi, and Frederic T Chong. Minimizing state preparations in variational quantum eigensolver by partitioning into commuting families. *arXiv preprint arXiv:1907.13623*, 2019.

[111] Hsin-Yuan Huang, Richard Kueng, and John Preskill. Predicting many properties of a quantum system from very few measurements. *Nature Physics*, 16(10):1050–1057, 2020.

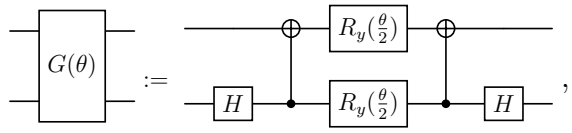
[112] Or Sharir, Amnon Shashua, and Giuseppe Carleo. Neural tensor contractions and the expressive power of deep neural quantum states. *Physical Review B*, 106(20):205136, 2022.

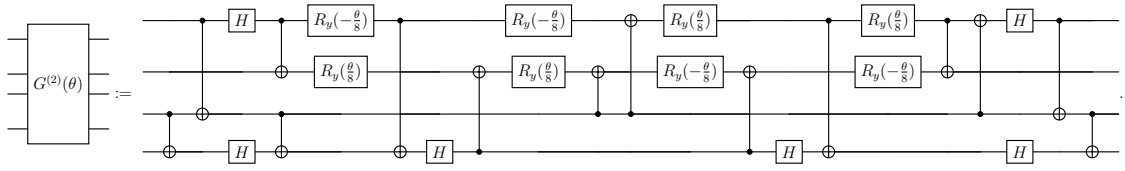
A

Appendix

A.1 Decompositions of Givens gates

The decompositions of $G(\theta)$ and $G^{(2)}(\theta)$ are as follows:





For an optimal scheme with fewer two-qubit gates, 2 CNOT gates are used in $G(\theta)$, while 14 CNOT gates are employed in $G^{(2)}(\theta)$.

A.2 Convergence evaluation

MR-REM utilizes improved multi-reference initial points, resulting in fewer iterations required for convergence in the VQE algorithm, as illustrated in several examples in Fig A.1.

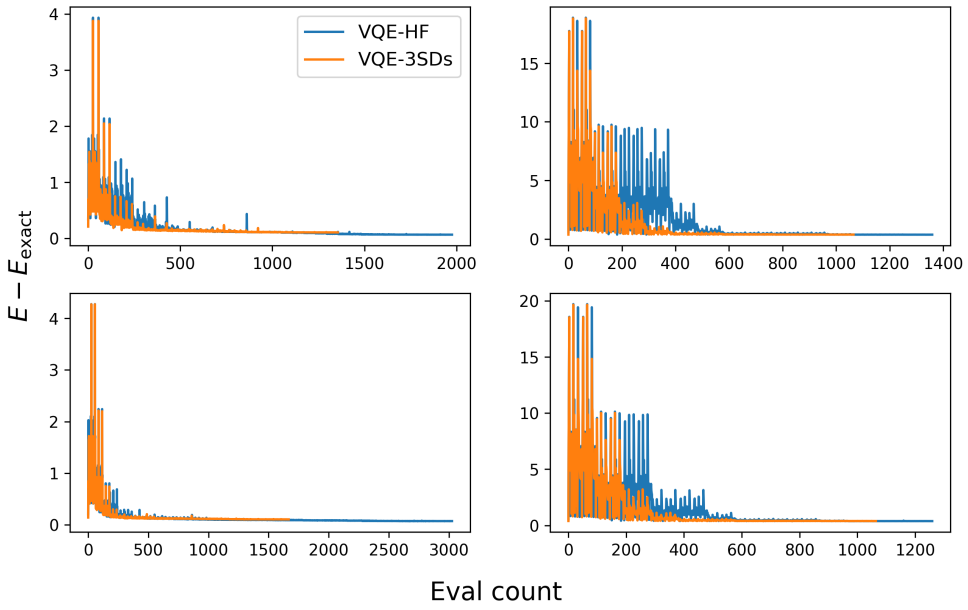


Figure A.1: Convergence iterations of VQE algorithm under the IMFIL optimizer for four example systems. Left: $H_2O(4e, 4o)$ at R = 2.45 Å (top) and R = 1.85 Å(bottom), and Right: $F_2(10e, 6o)$ at R = 2.7 Å (top) and R = 2.1 Å (bottom).

DEPARTMENT OF SOME SUBJECT OR TECHNOLOGY
CHALMERS UNIVERSITY OF TECHNOLOGY
Gothenburg, Sweden
www.chalmers.se

CHALMERS
UNIVERSITY OF TECHNOLOGY