



UNIVERSITÀ
DEGLI STUDI
DELL'AQUILA

Dipartimento di Ingegneria e Scienze dell'Informazione e Matematica

Dottorato di Ricerca in Matematica e Modelli
XXXVII ciclo

Bridging quantum and classically computed quantum chemistry via information theory, application to circuit design.

SSD 03/CHEM-02

道生一，一生二，二生三，三生萬物

The Dao gives rise to one, one to two, two to three, three to everything

道德经 (Dao De Jing) , Laozi, 42

Preface

道生一

The perfection of the world around us, while strongly symptomatic of the complexity of its underlying parts, stand as evidence for the superiority of the quantum description against the classical one. While the naive interpretation of the second law of thermodynamics tells us that the entropy of the universe always increases, the quantum description guarantees that the entropy of a closed system maintains a pure characterization with null entropy. This on the other hand suggests that the classical point of view necessarily lacks part of the picture.

一生二

The description of part of the universe from a quantum perspective is however plagued by the many-body interaction, of which quantum chemistry is a dominant example. In this framework, the description of a non-isolated system cannot transcend from the description of its environment, and when the environment is also composed of other objects of interest (二生三), the accurate description of their interaction will result in an extremely difficult task.

三生萬物

The results of these complex interactions eventually converge to the realization of tangible chemical effects of all sorts, sparking in such manner the interest for the study of the correlations that made these effects possible in the first place.

These correlations are, in the end, the primary and main reason for the development of this thesis.

In the study of the interactions typical of quantum chemistry, quantum computing and information theory have only recently started to play some parts in the development of further theoretical understanding. Sparkled by the potential application of quantum computing algorithms to quantum chemistry, this encounter has up to now been mainly targeted for quantum chemistry on quantum computers, but has now taken the opposite direction, where typical quantum computing instruments can now be used to deepen the understanding of classically computed quantum chemical properties. This is the context of the development of the present thesis, in a hybrid terrain between quantum chemistry and quantum information theory.

In this context, a recurring goal is to retrieve groundstates of typical many-body Hamiltonians on quantum devices, in this thesis the task is assigned to the Variational Quantum Eigensolver (VQE), a prominent algorithm that has revived hope of the application of quantum computing in the context of Near Intermediate Scale Quantum (NISQ) devices. However, this task is often plagued by convergence problems and limited capabilities of the existing hardware. Part of this thesis is dedicated to improving over these limitations, by using pre-assumed knowledge regarding the expected groundstate. Following the stated principle of hybridization between quantum and

classically computed quantum chemistry, this knowledge is based on classically computed guess of the groundstate.

On the other hand, this knowledge whose use we plan for the VQE improvements is by itself not easily obtained, as handling of wavefunctions in their space of definition is typically of exponential cost. While this is usually done efficiently by Tensor Network (TN) based methods, in the specific case of quantum chemistry, the complex and spatially spread two-body interactions of the electronic Hamiltonian limits the possible applications to relatively small-sized systems. This motivates the research for methods able to treat these wavefunctions efficiently.

In the literature, simulations of both quantum and classically computed quantum chemistry have been proven to depend strongly on the chemical intuition encoded on the second quantized Hamiltonian through the choice of orbital basis. This degree of freedom, while partially a curse for the unavoidable approximation that a limited basis set brings, allows one to choose a specific nature of the retrievable wavefunction, which will then describe the quantity of interest for which the basis was chosen in the first place. An exemplary case is given by the Localized Orbitals (LO), which express correlations typical of the chemical intuition of classical chemistry. On the other hand, the Natural Orbitals (NO), with their intuitive description of the occupation of an orbital, have been proven to facilitate the retrieval and description of the wavefunction in their own basis, which is undoubtedly a useful property for all variational methods. Given these different properties, one could also wonder how to characterize wavefunctions obtained with different sets of orbitals in terms of complexity and expressed correlation patterns. This natural curiosity is also partially addressed by this thesis.

To summarize, this thesis mainly aims to provide a comprehensive understanding of how correlations in quantum chemistry can be handled, studied and used via quantum information theory. To do this we will first of all introduce all the necessary mathematical and methodological tools, this is done in the first part I, which is divided into two main chapters. The first Chapter (1) introduces the quantum information theoretical background at the base of the anticipated discussion, we take care of this task by explaining both the mathematical framework in Section 1.1 and by introducing the field of quantum processing 1.2.

We then repeat the same idea for the necessary tools regarding Quantum Chemistry 2, with a first part dedicated to the mathematical foundations of second quantization 2.1, we then explain how this mathematical framework can be mapped on quantum computers 2.2, to then conclude introducing the current classically computed quantum chemistry methods 2.3, most of which are used in the anticipated interconnection between quantum chemistry and quantum computing.

These chapters, overall do not involve any development exclusive of this thesis, if not personal reinterpretation and elaboration of existing material. This changes from II onwards, where we start displaying the advancement brought by this thesis. Please note that the chapters are not listed in chronological order of development, but rather in the most suiting order for presenting the overall thesis. We begin with the work of chapter 3, in which we show how we can use classically computed approximated groundstates to retrieve quantum information quantities. This capability, besides its extensive use in the following chapters, has the great advantage of allowing the treatment of numbers of qubits which is unfeasible for any other existing Tensor Network method. We continue with the first use made of this algorithm in chapter 4, where we analyze the differences in quantum and classical correlations spanning between Hartree-Fock Canonical Orbitals (HFCO) or Natural

Orbitals (NO). The results display a marked difference between these two kinds of orbitals, as while the HFCO show some level of quantum-only correlations, the NO are mainly characterized by classical correlations.

So far, the shown research has concerned only quantum information analysis of wavefunctions of classically computed quantum chemistry. But these sorts of studies can be proven to be useful for quantum computing as well. In chapter 5 we show a first implementation of this idea. In this work, we implement a hardware-efficient ansatz for the VQE algorithm that makes use of known correlations deduced from MP2 wavefunctions with the use of the 3 method. With the further use of the NO, this newly developed ansatz has shown statistical and resource-wise advantage compared to other commonly used hardware-efficient ansätze. This is not only true for quantum chemistry, as with chapter 6 we demonstrate that further development of this idea can be also applied to the Heisenberg spin model. In this specific case, given the efficiency of mapping the used hamiltonian to Matrix Product Operator (MPO), we used a DMRG calculation for the retrieval of the expected correlation pattern which is then used in the development of the ansatz. Also in this case, the use of tailored ansatz for the system has proved decisive for the obtaining of statistically and resource-wise favorable results compared to other ansätze. Furthermore, with the advancements of the algorithm to include many layers of entanglement, we can obtain objectively good results even leaving the comparison for resources aside.

Eventually, after having explained all these methods, we finalize this thesis with the concluding chapter.

Contents

List of Figures vii

List of Tables viii

Acronyms x

I Theory 1

1 Quantum Information 2
- 1.1 Quantum Information Theory . . . 2
  - 1.1.1 Tensor Network . . . 11
  - 1.1.2 Matrix Product States (MPS) . . . 12
  - 1.1.3 Matrix Product Operators (MPO) . . . 13
  - 1.1.4 Density Matrix Renormalization Group Algorithm with MPS . . . 13
- 1.2 Quantum Computing . . . 16
  - 1.2.1 Quantum gates . . . 16
  - 1.2.2 Quantum advantage and how to get there . . . 17
  - 1.2.3 VQE . . . 20

2 Quantum Chemistry 24
- 2.1 Second Quantization . . . 26
  - 2.1.1 Basis Set . . . 30
  - 2.1.2 Molecular Orbitals choice . . . 31
- 2.2 Fermionic-to-qubit Mapping . . . 34
  - 2.2.1 Motivation and examples of mappings . . . 35
- 2.3 Classical computational method . . . 38
  - 2.3.1 Hartree Fock . . . 38
  - 2.3.2 Möller-Plesset Perturbation Theory . . . 40
  - 2.3.3 Configuration Interaction . . . 41

II Developments 45

3 SparQ 46
- 3.1 Introduction . . . 46
- 3.2 Wavefunction representations and sparsity in Quantum Chemistry methods . . . 47
  - 3.2.1 Encoding of fermionic wavefunctions . . . 48
  - 3.2.2 Jordan-Wigner mapping . . . 49
- 3.3 Mapping generalization . . . 51
- 3.4 Efficient mapping permutation . . . 52
- 3.5 Particle-Hole duality . . . 53

3.6 SparQ . . . 54
3.6.1 Partial trace and observables . . . 55
3.6.2 Partial Trace . . . 55
3.6.3 Observables . . . 56
3.6.4 Direct tracing . . . 56
3.7 Computational details . . . 59
3.8 Results . . . 59
3.9 Conclusions . . . 63

4 Quantum Information reveals that orbital-wise correlation is essentially classical in Natural Orbitals 64
4.1 Methods and Theoretical Background . . . 65
4.1.1 Natural Orbitals . . . 65
4.1.2 Two-qubit and one-qubit density matrices . . . 66
4.1.3 Computational Details . . . 68
4.2 Results and Discussion . . . 71
4.2.1 Conclusions . . . 74

5 QIDA, Quantum Information Driven Ansatz 75
5.1 Method . . . 76
5.1.1 The scheme . . . 76
5.2 Computational details . . . 80
5.2.1 Tested Systems . . . 80
5.2.2 Construction of the Ansatz . . . 80
5.3 Results . . . 82
5.4 Conclusions . . . 87

6 Compact Multi-Threshold Quantum Information Driven Ansatz For Strongly Interactive Lattice Spin Models 88
6.1 Method . . . 89
6.1.1 Heisenberg Model . . . 89
6.1.2 SO(4) Gates . . . 90
6.2 Multi-QIDA . . . 91
6.2.1 Layers-Builder . . . 91
6.2.2 Circuit ansatz composition . . . 92
6.2.3 Iterative layered-VQE . . . 94
6.3 Computational details . . . 95
6.3.1 Spin Models . . . 95
6.3.2 Metrics and measure . . . 96
6.3.3 Simulation details . . . 97
6.4 3x4 Isotropic System Results . . . 97
6.4.1 QMI Map and Qubit-pairs selection . . . 98
6.4.2 Layer-building procedure . . . 99
6.4.3 Heuristic ansätze comparison . . . 99
6.4.4 Performance Analysis . . . 100
6.4.5 Convergence and precision . . . 101
6.4.6 Layer-wise VQE optimization routine . . . 102
6.4.7 Summary . . . 103
6.5 Conclusion . . . 103

Conclusion 104

List of Publication **106**

Bibliography **108**

List of Figures

1.1 QIT - Visual representation of a tensor network . . . 12
1.2 QIT - Visual representation of a MPS with three sites . . . 12
1.3 QIT - Visual representation of a MPO with three sites . . . 13
1.4 QIT - Division in subpart to calculate effective Hamiltonian . . . 14
1.5 Overview of the major players over different quantum computing platforms. Taken from [1] . . . 19
1.6 Quantum Computing - UCCSD Ansatz . . . 23
1.7 Quantum Computing - Heuristic Ansatz . . . 23

2.1 Quantum Chemistry - General ternary tree representation of fermionic-to-qubit mapping . . . 35
2.2 Quantum Chemistry - Ternary tree representation for famous mappings . . . 36

3.1 SparQ - Time requirement analysis for direct and quadratic trace . . . 60
3.2 SparQ - Mutual information for the water molecule at CISD level with Jordan-Wigner and Parity mapping . . . 61
3.3 SparQ - Total entropies as a function of the active space dimension and selection methods for a benzene molecule. . . . 62

4.1 Correlations in NO - Berillium Dymer, von Neumann and Shannon mutual information 69
4.2 Correlations in NO - Benzene Dymer, von Neumann and Shannon mutual information 70
4.3 Correlations in NO - COnvergence of INO procedure . . . 73

5.2 QIDA - Method summary representation . . . 76
5.1 QIDA - Mutual information matrices for tested systems . . . 77
5.3 QIDA - Cross entanglement emprical reduction representation . . . 79
5.4 QIDA - Empiricaly reduced maps . . . 79
5.5 QIDA - Represantation of ansatz from mutual information matrix . . . 81
5.6 QIDA - Permutation-wise VQE results . . . 82
5.7 QIDA - Violin plot of the results . . . 83
5.8 QIDA - Violin plot for the NH_3 molecule . . . 84
5.9 QIDA - Resource analysis . . . 85

6.1 Multi QIDA - $\mathbf{SO}(4)$ decomposition in simpler gates . . . 91
6.2 Multi QIDA - Pictorial representation of the workflows of Multi-QIDA . . . 91
6.3 Multi QIDA - V-shape i-th layer . . . 93
6.4 Multi QIDA - Additional layer for $\mathbf{SO}(4)$ gates . . . 93
6.5 Multi QIDA - QMI map of 3×4 Isotropic Heisenberg model . . . 98
6.6 Multi QIDA - 3×4 QMI pairings selection . . . 98
6.7 Multi QIDA - 3×4 model VQE results . . . 101
6.8 Multi QIDA - VQE optimization convergence path . . . 102

List of Tables

1.1 QIT - Clifford normalizing table 10

3.1 SparQ - Summary of tested molecular systems together with the computational details 59

4.1 Correlations in NO - Tested Systems and their computational details 67
4.2 Correlations in NO - Results of simulations 72

5.1 QIDA - Tested systems 80

6.1 Multi QIDA - Simulated systems 96
6.2 Multi QIDA - Experimental results for the 3×4 Isotropic Heisenberg model Hamiltonian divided by ansatz layout 100

Acronyms

AO Atomic Orbitals

CAS Complete Active Space

CC Coupled Cluster

CI Configuration Interaction

CISD CI with Singles and Doubles excitations

DM Density Matrix

DMRG Density Matrix Renormalization Group

FCI Full CI

HF Hartree-Fock

HFCO Hartree-Fock Canonical Orbitals

INO Iterative Natural Orbitals

JW Jordan-Wigner

LO Localized Orbitals

MO Molecular Orbitals

MP2 Møller-Plesset perturbation theory of second order

MPO Matrix Product Operator

MPS Matrix Product State

MRCI Multi-Reference Configuration Interaction

Multi-QIDA Multi-Threshold Quantum Information Driven Ansatz

NISQ Noisy Intermediate-Scale Quantum

NO Natural Orbitals

NOON Natural Orbitals Occupation Number

PQC Parameterized Quantum Circuit

QC Quantum Computing

QCh Quantum Chemistry

QIDA Quantum Information Driven Ansatz

SD Slater Determinants

SO Spin Orbitals

SparQ Sparse Quantum State Analysis

TN Tensor Network

TPCPM Trace Preserving Completely Positive Map

TT Ternary Tree

VQA Variational Quantum Algorithm

VQE Variational Quantum Eigensolver

Part I

Theory

Chapter 1

Quantum Information

Generally, any scientific process involves the acquisition and elaboration of information, this makes it necessary to study how such information is transferred and modified by the environment and by the scientific process itself. This is the most general task of information theory, which in its classical formulation, based on probability theory, has found great theoretical applications in communication and information technologies. On the other hand quantum mechanics, as one of the greatest scientific achievements of 20th century, has brought a revolutionary paradigm shift in recent science. This new paradigm, carried and exemplified by the concept of *wavefunction*, has been revealed to be a generalization of classical probability. It then results natural to extend classical information theory and processing to encompass the wavefunction formalism, giving rise to Quantum Information. The first section of this chapter is dedicated to theoretical development and the second one delves into quantum information technologies. Most of the information of this chapter comes from the text of Nielsen & Chuang [109] and the notes of the Quantum Information Theory course at ETH [2], together with the book by John Watrous [148].

However, considering the field's current transition into the recently introduced Noisy Intermediate-Scale Quantum (NISQ) era, this thesis emphasizes variational approaches, which stand out as some of the most promising approaches for today's quantum devices. The Parameterized Quantum Circuit (PQC) ansatz design is one of the key challenges in these algorithms, requiring careful consideration of both hardware limitations and the physical properties of the system being studied.

1.1 Quantum Information Theory

In quantum mechanics, the state of a system is represented by a wavefunction $\psi(x)$. For a single particle moving in one dimension, the wavefunction is a complex-valued function $\psi : \mathbb{R} \to \mathbb{C}$ such that

$$\psi(x) = \langle x|\psi\rangle, \tag{1.1}$$

where $|x\rangle$ is the position eigenstate corresponding to position x.

The wavefunction must satisfy the normalization condition:

$$\int_{-\infty}^{\infty} |\psi(x)|^2 \, dx = 1.$$

In the Dirac notation, the state ψ is represented by a ket $|\psi\rangle$ as in Eq. (1.1). The ket notation

provides a more abstract way of representing quantum states, which is particularly useful in dealing with more complex quantum systems.

In general, if the space of definition of the wavefunction is not continuous but discretely defined over the event space $\mathcal{X}$ called alphabet, the same notation can be retained by using the following form:

$$\sum_{x\in\mathcal{X}} |\psi(x)|^2 = \sum_{x\in\mathcal{X}} |\langle x|\psi\rangle|^2 = 1.$$

A ket $|\psi\rangle$ is an element of a Hilbert space $\mathcal{H}$, which is a complete[1] inner product space. The inner product in this space is denoted as $\langle\phi|\psi\rangle$, where $|\phi\rangle$ and $|\psi\rangle$ are kets.

For any two states $|\phi\rangle$ and $|\psi\rangle$, the inner product $\langle\phi|\psi\rangle$ is a complex number. The inner product has the following properties:

1. **Conjugate Symmetry**: $\langle\psi|\phi\rangle = \langle\phi|\psi\rangle^*$,
2. **Linearity**: $\langle\phi|(a|\psi_1\rangle + b|\psi_2\rangle) = a\langle\phi|\psi_1\rangle + b\langle\phi|\psi_2\rangle$,
3. **Positive-definiteness**: $\langle\psi|\psi\rangle \geq 0$ with equality if and only if $|\psi\rangle = 0$.

The projection operator $|\psi\rangle\langle\psi|$ is used to project onto the state $|\psi\rangle$. The completeness relation for a basis of states $|x\rangle$ in $\mathcal{H}$ is given by:

$$\sum_{x\in\mathcal{X}} |x\rangle\langle x| = I \tag{1.2}$$

where I is the identity operator. In the current mathematical framework is then unambiguous to write the wavefunction of the state ψ as:

$$|\psi\rangle = \sum_{x\in\mathcal{X}} (\langle x|\psi\rangle)\,|x\rangle = \sum_{x\in\mathcal{X}} c_x\,|x\rangle \tag{1.3}$$

Where now $c_x = \langle x|\psi\rangle$ is the decomposition of the wavefunction over an orthonormal basis $|x\rangle$.

We can now introduce the concept of Density Matrix (DM), which generalizes classical probability distributions.

Definition 1. *The density matrix ρ is an endomorphism defined on a Hilbert space $\mathcal{H}$ (End($\mathcal{H}$)) with the following properties:*

- Positive semi-definite, such that $||\rho v||_{\mathcal{H}} \geq 0 \quad \forall v \in \mathcal{H}$
- Unit trace; $tr(\rho) = \sum_{x\in\mathcal{X}} (\langle x|\,\rho\,|x\rangle) = 1$

For a quantum state $|\psi\rangle$, the density matrix is given by:

$$\rho = |\psi\rangle\langle\psi|. \tag{1.4}$$

States described by this equation are called *pure* states. However, the definition of DM is not exhausted by such kind of states, as the space of endomorphisms over $\mathcal{H}$ is convex, and thus any

[1] In the Cauchy sense

linear combination(1.5) with total weight one (to preserve trace unity) is again a valid DM. In the following, we will unambiguously write $\mathcal{H}$ to actually mean $End(\mathcal{H})$ as the context clarifies whether the object treated is a vector $v \in \mathcal{H}$ or a density matrix $\rho \in End(\mathcal{H})$.

$$\begin{aligned} \rho &= \sum_i c_i \rho_i \\ \sum_i c_i &= 1 \end{aligned} \tag{1.5}$$

DM of this form represent *mixed* states, where the term refers to the fact that such a combination can be used to represent also a classical ensemble of quantum states. The condition for discerning between the two situations is the quantity called $Purity = \mathrm{Tr}(\rho^2)$:

Theorem 1. *ρ represents a pure state $\longleftrightarrow$ $Tr(\rho^2) = 1$*

The opposite condition to a state being pure is the so-called *fully mixed* state. It is defined as the state over n qubits such that:

$$|i\rangle \langle j| = \delta_{i,j} \frac{1}{2^n}$$

It thus represents, both from the quantum and from the classical probability perspective, the state with maximal entropy.

Tensor product

In quantum mechanics and linear algebra, the tensor product is an operation that takes two vector spaces and produces another vector space. If $\mathcal{H}_A$ and $\mathcal{H}_B$ are two Hilbert spaces, their tensor product, denoted by $\mathcal{H}_A \otimes \mathcal{H}_B$, is a Hilbert space consisting of all linear combinations of elements of the form $|a\rangle \otimes |b\rangle$, where $|a\rangle \in \mathcal{H}_A$ and $|b\rangle \in \mathcal{H}_B$. This structure is useful for treating combined quantum systems in which the wavefunction is expressed in terms of both spaces. In the current general discussion, the basis $\mathcal{X}$ that we have used so far can represent any space that satisfies the requirements of a Hilbert space, however, one commonly used space is the space of the tensor product of n binary basis of $|0\rangle$ and $|1\rangle$ called qubit, so that any wavefunction in this space can be written as in Eq. (1.6). This specific case will be the topic of Sec. 1.2.

$$|\psi\rangle = \sum_{\mathbf{i} \in (0,1)^n} k_{\mathbf{i}} |\mathbf{i}\rangle \tag{1.6}$$

Partial Trace

Following a similar idea to how larger spaces are constructed from smaller ones via the tensor product, we can reduce the description of a space to its subcomponents through the operation known as the *partial trace*. This procedure reduces the dimensionality of the analyzed space by discarding information pertaining to the irrelevant subsystems, analogous to the marginalization process in classical probability theory:

$$\rho_A = \mathrm{tr}(\rho_{AB})_B = \sum_{b \in B} \langle b| \rho_{AB} |b\rangle \tag{1.7}$$

Where the b states represent an orthonormal basis of the space B.

Within this context of combined spaces, a state ρ_{AB} is called *separable* in A and B if it can be

written as:

$$\rho_{AB} = \sum_i p_i \rho_{i,A} \otimes \rho_{i,B} \tag{1.8}$$

For some $\rho_{i,A} \in \mathcal{H}_A$, $\rho_{i,B} \in \mathcal{H}_B$ and where again $\sum_i p_i = 1$. If a state cannot be written in such form then it is *entangled* in A and B. An example of such states are the Bell states, also called *maximally entangled states*, as is $|\psi\rangle = \frac{1}{\sqrt{2}}(|00\rangle + |11\rangle)$.

Once we introduced how to reduce the state to contributions of a precise subspace with Eq.1.7 we can define valuable tools to characterize them. The first one is the **Quantum Entropy** Eq.1.9 or *von Neumann* entropy, which measures the uncertainty or randomness of a quantum state, analogous to the Shannon entropy in classical information theory.

$$\begin{aligned} 0 \leq S(A,B) &= -\operatorname{tr}(\rho_{AB}\log(\rho_{AB})) \\ 0 \leq S(A) &= -\operatorname{tr}(\rho_A \log(\rho_A)) \\ \rho_A &= \operatorname{tr}_B(\rho_{AB}) \end{aligned} \tag{1.9}$$

In complete analogy with classical information theory, we can also define the **Quantum Mutual Information** (or von Neumann):

$$0 \leq I(A,B) = S(A) + S(B) - S(A,B) \tag{1.10}$$

The mutual information is an important tool to quantify the correlation of two quantum states defined over orthogonal subspaces, and while it has been long used as a metric for establishing entanglement between two subsystems, it is known to be quite a poor measure for it. As an example let us evaluate the mutual information of a Bell State such as we have seen before:

$$|\psi\rangle = \frac{1}{\sqrt{2}}\left(|00\rangle + |11\rangle\right).$$

For this state, the density matrix ρ_{AB} is:

$$\rho_{AB} = |\psi\rangle\langle\psi| = \frac{1}{2}\begin{pmatrix} 1 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 1 \end{pmatrix}.$$

The reduced density matrices for subsystems A and B are:

$$\rho_A = \operatorname{Tr}_B(\rho_{AB}) = \rho_B = \frac{1}{2}\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}.$$

The von Neumann entropy for each of these reduced density matrices is:

$$S(\rho_A) = S(\rho_B) = -\operatorname{Tr}(\rho_A \log \rho_A) = 1.$$

The von Neumann entropy for the composite system is: $S(\rho_{AB}) = 0$ since $|\psi\rangle$ is a pure state. The

mutual information is therefore:

$$I(A:B) = S(\rho_A) + S(\rho_B) - S(\rho_{AB}) = 1 + 1 - 0 = 2.$$

The clear fact that the mutual information is greater than any value reachable by probability distributions defined over the same alphabet (which would be 1) is the reason why it was long considered a good candidate for the estimation of entanglement of a state. The reason behind this misconception was the link between the clear non-separability of the bell states and at the same time the high quantum correlation shown by these states, which cannot be rivaled by classical probability distributions.

However, to show that non-separability does not imply that the mutual information must be higher than classical reachable values, let us now consider the following state:

$$\rho_W = p|\psi^-\rangle\langle\psi^-| + (1-p)\frac{I}{4},$$

where $|\psi^-\rangle$ is the (singlet) state:

$$|\psi^-\rangle = \frac{1}{\sqrt{2}}\left(|01\rangle - |10\rangle\right),$$

I is the identity matrix, and $0 \leq p \leq 1$ is a parameter that determines the mixture.

- When $p = 1$, the state is purely the singlet state, which is maximally entangled, thus $I(A, B) = 2$ as shown previously.
- For $0 < p < 1$, the state is a mixture of the singlet state and the maximally mixed state.
- When $p = 0$ the state is maximally mixed, with I=0.

By continuity, we can see that there is a whole range of values p for which the state is still entangled and yet it is unnoticed by the Quantum mutual information being smaller than 1.

For future use, we ought to define the Trace Preserving Completely Positive Map (TPCPM), which describe the mathematical behavior of operation that can be done on quantum states, thus it is the mathematical definition of the intuitive idea of a *quantum channel* [2, 148]:

Definition 2. *A Trace Preserving Completely Positive Map (TPCPM) is an operation O defined on space $\mathcal{H}_A$ such that*

$$\forall \quad \mathcal{H}_E, \rho_E \in \mathcal{H}_E, \quad \rho_A \in \mathcal{H}_A, \rho' = (\mathcal{I} \otimes O^\dagger)(\rho_E \otimes \rho_A)(\mathcal{I} \otimes O), \quad \rightarrow tr(\rho') = 1, \rho' \geq 0 \tag{1.11}$$

Measurement

A crucial tool in quantum mechanics, the measurement is the way through which quantum information is brought to classical information. For a state ρ and an orthonormal basis $\mathcal{X}$ of $\mathcal{H}$, is called measurement the quantum channel $\rho \rightarrow \mathcal{M}$ defined as:

$$\begin{aligned} P(x) &= \langle x| \rho |x\rangle \\ \mathcal{M}(\rho) &= \sum_{x \in \mathcal{X}} P(x) |x\rangle \langle x| \end{aligned} \tag{1.12}$$

Connection to classical probability

In the introduction of this chapter, we remarked how quantum information is a generalization of classical information. Naturally, the tools we have introduced so far must then become known classical tools if we consider classical probabilities in place of the quantum states we used so far. Let us begin by recalling the need for a space of events from the Hilbert space $\mathcal{H}$, taking without loss of generality the alphabet $\mathcal{X}$ of an orthonormal basis of $\mathcal{H}$. We can define a space of probability of $\mathcal{X}$. A probability distribution $P_{\mathcal{X}}(x)$ can be represented as a quantum state ρ_P in the $\mathcal{X}$ basis as:

$$\rho_{P_{\mathcal{X}}} = \sum_{x\in\mathcal{X}} P(x) \left|x\right\rangle \left\langle x\right| \tag{1.13}$$

We notice that, given that any orthonormal basis can be used, each one of them would define a different space of probability, but any density matrix constructed with Eq. (1.13) is unambiguously the representation of a classical state.

Now, a joint probability distribution $P_{A,B}(a,b)$ defined over two event spaces A and B can be represented as:

$$\rho_{P_{A,B}} = \sum_{a\in A}\sum_{b\in B} P(a,b) \left|a,b\right\rangle \left\langle a,b\right| = \sum_{a\in A}\sum_{b\in B} P(a,b)(\left|a\right\rangle\otimes\left|b\right\rangle)(\left\langle a\right|\otimes\left\langle b\right|) = \sum_{a\in A}\sum_{b\in B} P(a,b) \left|a\right\rangle \left\langle a\right|\otimes\left|b\right\rangle \left\langle b\right| \tag{1.14}$$

We can now notice two elements from (1.14). The first is that at the right-most hand side, we have the definition of a separable density matrix as in Eq. (1.8), thus obviously any classical state is not entangled. Second, taken the marginalization of $P_{A,B}(a,b)$ over B we have that:

$$\begin{aligned} P_A(a) &= \sum_{b\in B} P_{A,B}(a,b) \\ \mathrm{tr}(\rho_{P_{A,B}})_B &= \sum_{b'\in B} \left\langle b'\right| (\sum_{a\in A}\sum_{b\in B} P(a,b) \left|a\right\rangle \left\langle a\right| \otimes \left|b\right\rangle \left\langle b\right|) \left|b'\right\rangle = \\ &= \sum_{a\in A} \left|a\right\rangle \left\langle a\right| (\sum_{b'\in B}\sum_{b\in B} \delta_{b,b'} P(a,b)) = \sum_{a\in A} P(a) \left|a\right\rangle \left\langle a\right| \end{aligned} \tag{1.15}$$

Thus, when treating classical probability distributions, the partial trace operation is equal to the marginalization.

Let us now also give an analogous of the von Neumann measures. In classical information theory, we call entropy the following quantity:

$$\begin{aligned} S_{Sh}(A,B) &= -\sum_{a\in A, b\in A} P(a,b)\log(P(a,b)) \\ S_{Sh}(A) &= -\sum_{a\in A} P(a)\log(P(a)) \end{aligned} \tag{1.16}$$

Where the subscript stands for Shannon entropy[131], and $S_{Sh}(A,B)$ is for the joint spaces A and B. Now let us consider again the definition of quantum entropy of Eq. (1.9), and furthermore let us take the density matrix $\rho_{P_{A,B}}$ relative to the distribution $P_{A,B}(a,b)$ as in Eq. (1.14), if we realize that this density matrix is a diagonal matrix, then we have that:

$$
\begin{aligned}
&S(A,B) = -\operatorname{tr}(\rho_{P_{A,B}} \log(\rho_{P_{A,B}})) = \\
&\sum_{a' \in A} \sum_{b' \in B} \langle a' | \langle b' | (\sum_{a \in A} \sum_{b \in B} P_{A,B}(a,b) log(P(a,b)) |a\rangle \langle a| \otimes |b\rangle \langle b|) |a'\rangle |b'\rangle = \\
&\sum_{a' \in A} \sum_{b' \in B} \sum_{a \in A} \sum_{b \in B} P_{A,B}(a,b) log(P(a,b)) \delta_{a,a'} \delta_{b,b'} = S_{Sh}(A,B)
\end{aligned}
\tag{1.17}
$$

Thus for a classical probability distribution, the Shannon entropy and the von Neumann one coincide. Clearly, this also extends to the mutual information, where we have that:

$$I_{Sh}(A,B) = S_{Sh}(A) + S_{Sh}(B) - S_{Sh}(A,B) \tag{1.18}$$

and thus for classical states

$$I_{Sh}(A,B) = S_{Sh}(A) + S_{Sh}(B) - S_{Sh}(A,B) = S(A) + S(B) - S(A,B) = I(A,B)$$

.

We now see why mutual information is not a precise tool for analyzing entanglement in quantum states. Since it generalizes the Shannon mutual information, it contains contributions of both classical and quantum correlations. Furthermore, while one can extract the classical-only correlation form a state, one cannot be sure to isolate only the quantum correlations with the mutual information. Let us explain this concept more in detail. Suppose we have a general quantum state $\rho_{A,B}$, defined over space A and B, we can analyze this state employing the quantum mutual information $I(A,B)$, which will give an estimate of all the correlations in the two parts. However, to identify *some* classical correlations, one can try to reduce the state to a classical distribution by measuring the state in a computational basis of choice, say $\mathcal{X}$. Thus evolving the state through a projective measurement channel $\mathcal{M}$ as in Eq. (1.12), the resulting density matrix will then assume a diagonal form in the computational basis in which the measurement was undergone, just as in Eq. (1.14). As we showed before, for such measured state $I = I_{Sh}$, however, the original state was not $\mathcal{M}(\rho)$ but rather ρ, and thus we are interested in a relation between the mutual information I_ρ of the state ρ and how it compares with $I_{Sh,\mathcal{M}(\rho)}$, as this latter quantities represents the classical correlations captured by the specific computational base $\mathcal{X}$.

To establish this relation we start from Eq. (1.12) and given a TPCPM, indicated with $\mathcal{O}$, the following can be proven [148]:

Theorem 2. *Given a $\rho_{AB} \in \mathcal{H}_A \otimes \mathcal{H}_B$ and a TPCPM $\mathcal{O}$ acting only on space B, then:*

$$I_\rho(A,B) \geq I_{\mathcal{O}(\rho_{AB})}(A,B) \tag{1.19}$$

This relation states the intuitive fact that a local operation only on space B cannot increase the mutual information between A and B. In the case under analysis, a measurement can be considered as a TPCPM, since it is both trace-preserving and it doesn't alter the positivity of the density matrix.

Proof. For a measurement applied on one of the two subspaces, e.g. at space A ($\mathcal{M}(\rho)_A$), Eq. (1.19)

then becomes

$$I(A,B) \geq I_{\mathcal{M}(\rho)_A}(A,B) \tag{1.20}$$

Which repeated to the other subspace leads to:

$$\begin{aligned} I_{\mathcal{M}(\rho)_A}(A,B) &\geq I_{\mathcal{M}(\rho)_{A,B}}(A,B) \\ &= I_{\mathcal{M}(\rho)}(A,B) = I_{Sh}(A,B) \end{aligned} \tag{1.21}$$

where $\mathcal{M}(\rho)$ is the measurement over the whole space. The last equality of (1.21) due to the relations (1.17)(1.12). This finally leads us to:

$$I(A,B) \geq I_{Sh}(A,B) \tag{1.22}$$

□

This inequality states that given any bipartite (or multi-partite) state, the total (quantum and classical) correlation between two different parts is going to be greater than the classical-only correlation. This is going to be used in the following works 4.

Stabilizer formalism

The stabilizer formalism is a powerful and elegant method that, using the language of group theory, describes a large class of quantum error-correcting codes, known as stabilizer codes. In our context, we introduce the stabilizer formalism to arrive at another notion of complexity in quantum states so that they can be simulated efficiently.

The stabilizer group $\mathcal{S}$ of is an abelian subgroup of the Pauli group $\mathcal{P}_n$ on n qubits, excluding $-I$. The Pauli group $\mathcal{P}_n$ consists of all n-fold tensor products of the Pauli matrices $\{X, Y, Z\}$ together with the identity matrix I, with overall phase factors $\{\pm 1, \pm i\}$.
The stabilizer group $\mathcal{S}$ defines the stabilized space $\mathcal{C}$ as the common $+1$ eigenspace of all elements in $\mathcal{S}$:

$$\mathcal{C} = \{|\psi\rangle \in \mathcal{H}_2^n : M|\psi\rangle = |\psi\rangle, \forall M \in \mathcal{S}\}.$$

A stabilizer space is specified by its stabilizer group $\mathcal{S}$, and it is the simultaneous $+1$ eigenspace of all stabilizer generators. If the stabilizer group $\mathcal{S}$ has $n-k$ independent generators, the stabilized space $\mathcal{C}$ will be a subspace of dimension 2^k.

Let us now consider the case $k = 0$, then the stabilized space $\mathcal{C}$ will identify a single state $|\mathcal{C}\rangle$ of the space $\mathcal{H}$. We call *stabilizer state* any such state $|\mathcal{C}\rangle$ for which there can be found n commuting Pauli operators for which $|\mathcal{C}\rangle$ is a $+1$ eigenvector.

Clifford operations

The Clifford unitaries are the set $\mathcal{N}$ of unitaries $U \in \mathcal{U}$ of the space $\mathcal{H}$ that normalize the Pauli group.

$$\mathcal{N} = \{U \in \mathcal{U} : U\sigma U^\dagger \in \mathcal{P}_n, \forall \sigma \in \mathcal{P}_n\} \tag{1.23}$$

It can be easily proven[109] that this set forms a group, and its generators are made of all n-fold tensor products of the following unitaries:

- The Hadamard gate; $H = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix}$
- The Phase gate; $S = \begin{pmatrix} 1 & 0 \\ 0 & i \end{pmatrix}$
- The CNOT gate; $\text{CNOT} = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 \end{pmatrix}$

One obvious consequence of this definition is that, since the Pauli matrices normalize themselves, then $\mathcal{P}_n \subsetneq \mathcal{N}$.

$U \in \mathcal{N}$	**Input**	**Output**
$CNOT$	X_1	X_1X_2
	X_2	X_2
	Z_1	Z_1
	Z_2	Z_1Z_2
H	X	Z
	Z	X
S	X	Y
	Z	Z
X	X	X
	Z	$-Z$
Y	X	$-X$
	Z	$-Z$
Z	X	$-X$
	Z	Z

Table 1.1: Table representing the effect of Clifford operations on the Pauli group. Figure taken from [109]

We can now study what happens to a stabilizer set $\mathcal{S}$ if we apply a Clifford unitary onto its stabilized state $|\mathcal{C}\rangle$.

$$\left|\mathcal{C}'\right\rangle = U\left|\mathcal{C}\right\rangle = UM\left|\mathcal{C}\right\rangle = UMU^\dagger U\left|\mathcal{C}\right\rangle = M'U\left|\mathcal{C}\right\rangle = M'\left|\mathcal{C}'\right\rangle \tag{1.24}$$

Thus, the new stabilizer state $U|\mathcal{C}\rangle = \mathcal{C}'$ is stabilized by the new stabilizer group

$$\mathcal{S}' = \{UMU^\dagger, \forall M \in \mathcal{S}\} \tag{1.25}$$

Given the overall structure, it is then clear that to follow a Clifford evolution of a stabilizer state $\mathcal{C}$ to $\mathcal{C}'$ we can simply follow the evolution of its stabilizing group from $\mathcal{C}$ to $\mathcal{C}'$ in an Heisemberg operatorial manner. Furthermore, we also notice that normalizing evolution of the Clifford unitaries can be calculated efficiently through tabulated values such as the one in Tab.1.1. We have then shown how we can efficiently follow the evolution of a state as long as this evolution is only composed of Clifford unitaries. This is the so-called *Gottesman-Knill* theorem.[4]

The interesting point of this approach is that entanglement, which has long been thought to be the key resource for the advantage of quantum computing, can be easily generated using Clifford

unitaries. This is shown by the example of the maximally entangled state $\frac{1}{\sqrt{2}}(|00\rangle + |11\rangle)$, which can be generated from the state $|00\rangle$ with an H gate and a CNOT gate, both Clifford gates. As a generalization to the Gottesmann-Knill theorem, other methods focusing on *near*-Clifford unitaries can also efficiently be simulated [33], although with an exponential cost in terms of the number of non-clifford operations. This can be done by following the non-Clifford evolution of a single stabilizer state $|\mathcal{C}\rangle$ into wavefunctions of the form of Eq. (1.26), where each $|\mathcal{C}_i\rangle$ is a stabilizer state relative to a different stabilizer group $\mathcal{S}_i$

$$U|\mathcal{C}\rangle = \sum_i a_i |\mathcal{C}_i\rangle \qquad \text{with} \quad U \notin \mathcal{N}, \qquad \sum_i a_i^2 = 1 \tag{1.26}$$

Overall, this framework paints a picture in which entanglement, while surely important for resource analysis, is not a sufficient condition for the characterization of quantum-only complexity. For this reason, one quantity thought to properly represent the complexity of a wavefunction is now believed to be the number of stabilizer states composing a wavefunction of the kind as in Eq. (1.26).

1.1.1 Tensor Network

This section introduces Tensor Network (TN), a powerful and versatile tool for representing variational wavefunctions and other high-dimensional data structures. This section is based on the more complete overview by Schwollock[130].

A *tensor* is a generalization of scalars, vectors, and matrices to higher dimensions. Formally, a tensor can be viewed as a multidimensional array of complex or real numbers, with the number of indices (or "legs") indicating its rank or dimensionality. For example, a scalar is a rank-0 tensor, a vector is a rank-1 tensor, and a matrix is a rank-2 tensor. A tensor of rank-3 $T^{\sigma_1,\sigma_2,\sigma_3}$ has three indices, each of which runs over some set of possible values. Mathematically, a tensor can be represented as:

$$T^{\sigma_1,\sigma_2,\ldots,\sigma_N} \in \mathbb{C}^{d_1 \times d_2 \times \cdots \times d_N}$$

where d_i is the dimension of the ith index. The TN approach builds upon this concept by representing large and complex tensors as networks of smaller interconnected tensors, which can then be contracted efficiently. A tensor as $T^{\sigma_1,\sigma_2,\ldots,\sigma_N} \in \mathbb{C}^{d_1 \times d_2 \times \cdots \times d_N}$ can be represented in tensor network formalism by also being expressed in terms of subcomponents, as an example, let us take the following tensor:

$$T_{\sigma_1\sigma_2\sigma_3\sigma_4\sigma_5} = \sum_{a_1,a_2,a_3,a_4,a_5} A^{[1]}_{\sigma_1,a_1,a_2} A^{[2]}_{\sigma_2,a_2,a_3} A^{[3]}_{\sigma_3,a_3,a_4} A^{[4]}_{\sigma_4,a_4,a_5} A^{[5]}_{\sigma_5,a_5,a_1} \tag{1.27}$$

In this equation, the tensor $T_{\sigma_1\sigma_2\sigma_3\sigma_4\sigma_5}$ is expressed through tensor contraction of other 3-rank tensors, the contracted indices are called *virtual* indices while the remaining are called *physical* indices. An illustration representing this tensor through tensor network graphs is shown in Fig. 1.1.

Originally developed for classical simulations of statistical models and many-body quantum systems, the TN method approximates the wavefunction of a quantum system by decomposing the large tensor representing the full wavefunction into more manageable smaller tensors. This decomposition reduces the computational complexity and memory requirements, making TN a highly

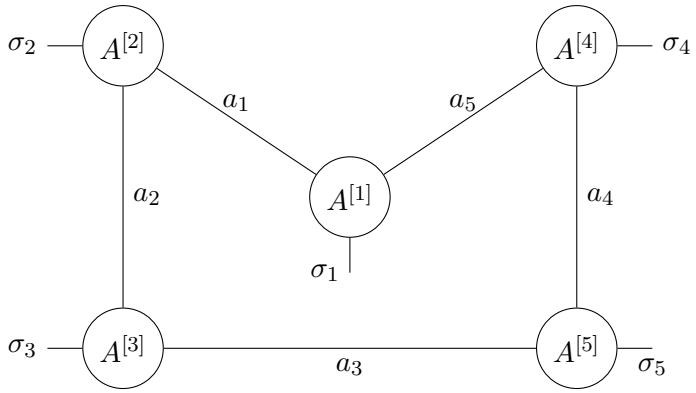


Figure 1.1: Visual representation of the tensor network represented by Eq. (1.27)

efficient framework for representing quantum systems, especially in scenarios where the wavefunction is sparse. Beyond physics and quantum chemistry, TNs have also found applications in fields like data analysis and machine learning. Given their flexibility and computational advantages, TNs will serve as the primary tool for comparison in the later sections of this thesis 3.

1.1.2 Matrix Product States (MPS)

A specific type of TN that is widely used in quantum systems is the Matrix Product State (MPS). An MPS represents a quantum wavefunction by linearly decomposing the state into a chain of interconnected tensors. In an MPS, the wavefunction of a quantum system with N sites can be expressed as a product of matrices (or tensors) where each tensor corresponds to one site. The wavefunction $|\psi\rangle$ can be written as:

$$|\psi\rangle = \sum_{\sigma_1,\cdots\sigma_N} \sum_{a_0\cdots,a_N} A^{[1]}_{\sigma_1,a_1} A^{[2]}_{\sigma_2,a_1,a_2} \dots A^{[N]}_{\sigma_N,a_{N-1}} |\sigma_1,\sigma_2,\dots,\sigma_N\rangle \tag{1.28}$$

where $A^{[k]}_{\sigma_k,a_{k-1},a_k}$ is a tensor with a physical index σ_k, representing the local degrees of freedom, and up to two virtual indices connecting adjacent tensors. The virtual indices serve as a means of encoding correlations between the sites, and the maximum size of these indices is controlled by a parameter χ, known as the *bond dimension*. A visual representation of an MPS with three sites is given by Fig. 1.2.

The expressibility of an MPS is determined by the bond dimension χ. For small χ, the MPS can efficiently represent states with limited entanglement, but as χ increases, the MPS can approximate more complex and entangled states. When χ is large enough, an MPS can exactly represent any quantum state.

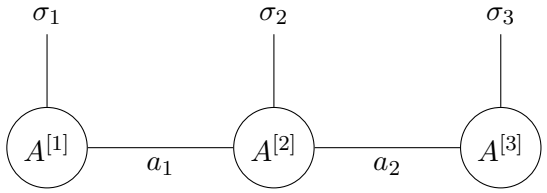


Figure 1.2: Visual representation of a MPS with three sites

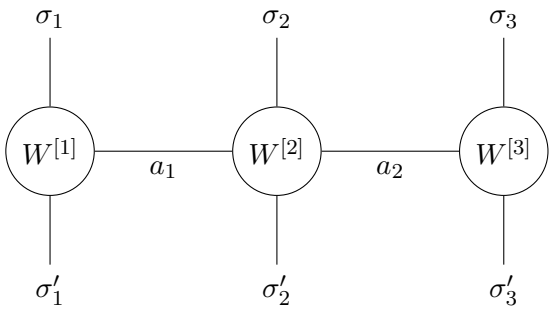


Figure 1.3: Visual representation of a MPO with three sites

1.1.3 Matrix Product Operators (MPO)

To describe physical systems, we also need to represent quantum operators, such as the Hamiltonian H, which defines the system's energy. In the TN framework, operators can be represented using Matrix Product Operator (MPO). An MPO is the operator analog of the MPS and is similarly expressed as a chain of matrices or tensors. Just as the wavefunction can be decomposed into a product of tensors, the operator can be decomposed into a product of tensors representing local operators acting on each site.

An MPO for a Hamiltonian H acting on N sites is written as:

$$H = \sum_{\sigma_1,\sigma'_1\cdots\sigma_N,\sigma'_N} \sum_{a_0\cdots,a_N} W^{[1]}_{\sigma_1\sigma'_1,a_1} W^{[2]}_{\sigma_2\sigma'_2,a_1a_2} \cdots W^{[N]}_{\sigma_N\sigma'_N,a_{N-1}} \tag{1.29}$$

Here, $W^{[k]}_{\sigma_k\sigma'_k,a_{k-1},a_k}$ represents the local operator acting on the k-th site, with virtual indices a_{k-1}, a_k connecting the operators at adjacent sites. A visual representation of an MPO with three sites is given by Fig. 1.3. The bond dimension of the MPO similarly controls the complexity and accuracy of the operator representation.

Often the kind of MPO needed is the one of an Hamiltonian which needs to be minimized by an MPS. To this scope, different approaches can be taken depending on the kind of problem that is being solved, and for the specific case of the electronic Hamiltonian introduced in the following chapter 2 we refer to the proper work by Keller *et al* for efficient implementation [77].

1.1.4 Density Matrix Renormalization Group Algorithm with MPS

In this work, we use the Density Matrix Renormalization Group (DMRG) algorithm [129, 130, 149, 150] in conjunction with the MPS representation. DMRG is a variational algorithm that optimizes an MPS wavefunction, typically aiming to find the ground state of a Hamiltonian H. It involves the following steps:

The DMRG algorithm consists of three main steps:

1. Initialization
2. Sweep procedure
3. Measurement and convergence check

Let us now break down each step in detail.

Initialization

Start with an initial Matrix Product State (MPS) $|\psi\rangle$ and Matrix Product Operator (MPO) representation of the Hamiltonian H.

$$|\psi\rangle = \sum_{\sigma_1,\ldots,\sigma_N} A_{\sigma_1} A_{\sigma_2} \ldots A_{\sigma_N} |\sigma_1, \sigma_2, \ldots, \sigma_N\rangle \tag{1.30}$$

Where A_{σ_i} are site tensors, and σ_i are physical indices. Differently than to (1.28), we have now suppressed the dependency over the virtual indices as they will always be summed over.

Sweep Procedure

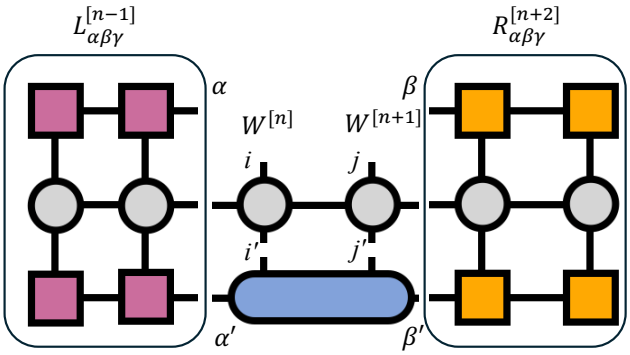


Figure 1.4: Division in subpart to calculate effective Hamiltonian as in Eq.(1.31), indexes across the *legs* as in Eq. (1.35)

Left-to-right sweep

For each pair of sites $(n, n+1)$:

1. Construct the effective Hamiltonian:

$$H_{\text{eff}} = L^{[n-1]} \cdot W^{[n]} \cdot W^{[n+1]} \cdot R^{[n+2]} \tag{1.31}$$

Where $L^{[n-1]}$ and $R^{[n+2]}$ are left and right environments:

$$L^{[n-1]}_{\alpha,\beta,\gamma} = \sum_{\sigma_1,\ldots,\sigma_{n-1}} (A_{\sigma_1} \ldots A_{\sigma_{n-1}})^*_\alpha H_{\beta,\gamma} (A_{\sigma_1} \ldots A_{\sigma_{n-1}})_\alpha \tag{1.32}$$

$$R^{[n+2]}_{\alpha,\beta,\gamma} = \sum_{\sigma_{n+2},\ldots,\sigma_N} (A_{\sigma_{n+2}} \ldots A_{\sigma_N})_\alpha H_{\beta,\gamma} (A_{\sigma_{n+2}} \ldots A_{\sigma_N})^*_\alpha \tag{1.33}$$

Here $H_{\beta,\gamma}$ still contains a summation across all σ indices.

2. Solve the local eigenvalue problem:

$$H_{\text{eff}}|\psi\rangle = E|\psi\rangle \tag{1.34}$$

Explicitly:

$$\sum_{\alpha'\beta'i'j'} H^{\text{eff}}_{\alpha\beta ij,\alpha'\beta'i'j'}\Psi_{\alpha'\beta'i'j'} = E\Psi_{\alpha\beta ij} \tag{1.35}$$

3. Perform Singular Value Decomposition (SVD) on the optimized two-site tensor:

$$\Psi_{\alpha\beta ij} = \sum_{\gamma} U_{\alpha i\gamma} S_\gamma V^\dagger_{\gamma\beta j} \tag{1.36}$$

4. Truncate singular values:

$$S_\gamma \to S'_\gamma \text{ for } \gamma \leq \chi \tag{1.37}$$

Truncation error:

$$\varepsilon = 1 - \sum_{\gamma=1}^{\chi} (S_\gamma)^2 \tag{1.38}$$

5. Update MPS tensors:

$$A^{[n]}_{\alpha i} = U_{\alpha i\gamma}\sqrt{S'_\gamma} \tag{1.39}$$

$$A^{[n+1]}_{\beta j} = \sqrt{S'_\gamma} V^\dagger_{\gamma\beta j} \tag{1.40}$$

6. Move to the next pair of sites

It has to be noted that during sweeps one has to maintain a mixed canonical form:

$$\sum_i A^{[n]}_{\alpha i}(A^{[n]}_{\beta i})^* = \delta_{\alpha\beta} \quad \text{(Left-orthogonalization)} \tag{1.41}$$

$$\sum_i (A^{[n]}_{\alpha i})^* A^{[n]}_{\beta i} = \delta_{\alpha\beta} \quad \text{(Right-orthogonalization)} \tag{1.42}$$

Right-to-left sweep

Repeat the process in the opposite direction.

Measurement and Convergence Check

1. Calculate energy expectation value:

$$E = \frac{\langle\psi|H|\psi\rangle}{\langle\psi|\psi\rangle} \tag{1.43}$$

Explicitly:

$$E = \frac{\sum_{\sigma_1,\ldots,\sigma_N,\sigma'_1,\ldots,\sigma'_N} (A_{\sigma_1}\ldots A_{\sigma_N})^* H_{\sigma_1\ldots\sigma_N,\sigma'_1\ldots\sigma'_N} (A_{\sigma'_1}\ldots A_{\sigma'_N})}{\sum_{\sigma_1,\ldots,\sigma_N} |(A_{\sigma_1}\ldots A_{\sigma_N})|^2} \tag{1.44}$$

2. Check convergence:

$$|E_n - E_{n-1}| < \varepsilon \tag{1.45}$$

where ε is a small threshold.

The overall computational time complexity of DMRG scales as $\mathcal{O}(ND\chi^3)$, where N is the number of sites, D is the number of degrees of freedom per site, and χ is the bond dimension. Justified by this computational cost, the limitations of DMRG strictly depend on the bond dimension. For this reason, it is usually difficult to give a limitation in terms of the system-size, it is clear, however, that the size will influence the hardness of the problem via the size of the spanned Hilbert space and the minimum required bond dimension. In general, fixing the precision of the simulation in terms of the energy will result in a minimum required $\chi_{min}(N)$ that will depend on the system size.

Finding this dependency is not trivial, as it also strongly depends on the kind of system studied. For instance, in QCh, our main field of interest for this thesis, the time and memory scaling of the algorithm is recognized as $\mathcal{O}(\chi^3 N^4)$ and $\mathcal{O}(\chi^2 N^2)$ respectively [78]. One can then see how any polynomial dependency of χ in N will get the computational cost to a worse level than other P-HF methods. For all practical purposes, DMRG calculations for quantum chemistry rarely exceed a few dozen orbitals, and when they do, they require massive HPC infrastructure [100].

1.2 Quantum Computing

While we have so far treated quantum spaces of general dimension in quantum computing analyses typically occur in a Hilbert space $\mathcal{H}$ consisting of multiple separable two-dimensional spaces known as *qubits* in tensor product within each other. The entire space of N qubits $\mathcal{H}$ is then

$$\mathcal{H} = \bigotimes_{i}^{N} \mathcal{H}_i \tag{1.46}$$

with a total dimension of $\mathcal{H}$ is 2^N [109]. The reason for this specific structure of a *two-level* system in place of other possibilities is due to the classical counterpart of information theory and processing, based on classical *bits*.

1.2.1 Quantum gates

In quantum computing, quantum gates are the fundamental building blocks for quantum circuits. Their role is to manipulate qubits to perform quantum computations. Each quantum gate performs a specific unitary transformation on the state of one or more qubits.

Single-qubit gates act on individual qubits and are indicated as $\boxed{U}$, where U is replaced with the respective name. Some common single-qubit gates include: **Pauli Gate** Analogous to the famous matrix operators these gates are:

$$X = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \qquad Y = \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix} \qquad Z = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}$$

Hadamard Gate (H)

This gate creates superposed states from basis states:

$$H = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix}$$

Phase Gate (S)

This gate applies a phase shift of $\pi/2$:

$$S = \begin{pmatrix} 1 & 0 \\ 0 & i \end{pmatrix}$$

All these matrices are defined as operators over the basis $|0\rangle$, $|1\rangle$ (in this order) of a single qubit.

Multi-qubit gates act on two or more qubits and are essential for creating entanglement. One of the most important multi-qubit gates is the CNOT (controlled-NOT) gate.

$$\text{CNOT} = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 \end{pmatrix} \qquad \text{} \tag{1.47}$$

While these gates are useful quantum computing tools, widely necessary for computing, they are not universal, meaning that they cannot implement all the unitary operations over the qubits. To do that, one can either add in the set of available gates one of the single qubit arbitrary rotations:

$$R_X(\theta) = \begin{pmatrix} \cos(\theta/2) & -i\sin(\theta/2) \\ -i\sin(\theta/2) & \cos(\theta/2) \end{pmatrix} \qquad R_Y(\theta) = \begin{pmatrix} \cos(\theta/2) & -\sin(\theta/2) \\ \sin(\theta/2) & \cos(\theta/2) \end{pmatrix} \qquad R_Z(\theta) = \begin{pmatrix} e^{-i\theta/2} & 0 \\ 0 & e^{i\theta/2} \end{pmatrix}$$

the T gate $T = \begin{pmatrix} 1 & 0 \\ 0 & e^{i\pi/4} \end{pmatrix}$, or a three-qubit gate called Toffoli gate (controlled-controlled-NOT)

$$\text{Toffoli} = \begin{pmatrix} 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 \end{pmatrix} \qquad \text{} \tag{1.48}$$

In any case, the aim is to extend the Clifford set to a universal set of gates, so that any unitary can be implemented.

The last crucial missing operation used in quantum computing[2] is the measurement, indicated with  and implementing the channel of (1.12) for a single qubit.

1.2.2 Quantum advantage and how to get there

Quantum computing is expected to revolutionize various fields, leveraging the principles of quantum information introduced so far, it promises to provide computational power beyond that of classical computers. One significant expected advantage of quantum computing, and the main focus of this

[2]There would also be conditioned operations but are not commonly used in near-term quantum computing

work, is its ability to handle tasks that require simulating quantum systems[13, 99, 113], ideally breaking the exponential scaling of full treatment of these tasks, albeit with noticeable disagreement in literature with regards to the actual scaling expectations [88]. One more impactful application is thought to be cryptography, as quantum algorithms like Shor's algorithm [132] can factor large numbers exponentially faster than classical algorithms, posing a threat to current encryption methods and giving rise to the field of Post-quantum criptography [22]. Other relevant applications are represented by the fields of logistics and supply chain management, together with finance and portfolio management. [52] Overall, while promising algorithms exist, it is important to remember that not all problems benefit from quantum computing, and ongoing research is needed to determine which problems are best suited.

Despite these promising prospects, quantum computing faces significant challenges. Quantum computers are highly susceptible to errors caused by decoherence and quantum noise. [109] Effective quantum error correction requires a large overhead of physical qubits, which is currently only feasible for demonstrative purposes with current technology. This is because scalability is another challenge, as building and maintaining a large number of high-quality qubits is technically demanding.

Finally, the cost and infrastructure required for quantum computing represent significant barriers. Quantum hardware is currently expensive to develop and maintain, requiring highly controlled environments.

All these limitations are being addressed by trying many different hardware platforms for the implementation of the quantum devices, each of them with different weaknesses and strengths:

- **Superconducting Qubits:** These qubits use superconducting circuits cooled to extremely low temperatures to create and manipulate quantum states. The main advantage for adoption at the basis of superconducting qubits is the similarity with silicon classical processors in terms of ideal production and small dimension of the chips, together with fast circuit execution. However, controlling the qubits while also increasing their number has remained a challenge so far. Companies like IBM, Google, IQM and Rigetti are leading in this technology.

- **Trapped Ions:** Ions trapped in electromagnetic fields are used as qubits, manipulated using laser beams. IonQ and Quantinuum are prominent players in this technology. Trapped ion systems are known for their high-fidelity operations and long coherence times. This features are counterbalanced by the also long gate duration, especially for multiple-qubits operations, which often requires addressing and reordering of the qubit lattice.

- **Rydberg Atoms:** Rydberg atoms use highly excited states of atoms to form qubits, which can be manipulated with laser light. They offer strong interactions over relatively long distances, making them suitable for entanglement and quantum gate operations. Companies like Pasqal are exploring this technology. While gaining much popularity in recent years, the main drawback is again the duration of execution of the gates together with the fact that neighboring qubits can't be addressed independently at the same time due to the blockade effect.

- **Cold Atoms:** Neutral atoms are cooled and trapped in optical lattices or tweezers, creating qubits that can be manipulated with laser light. Cold atom systems benefit from the scalability and long coherence times. Researchers and companies like ColdQuanta are developing

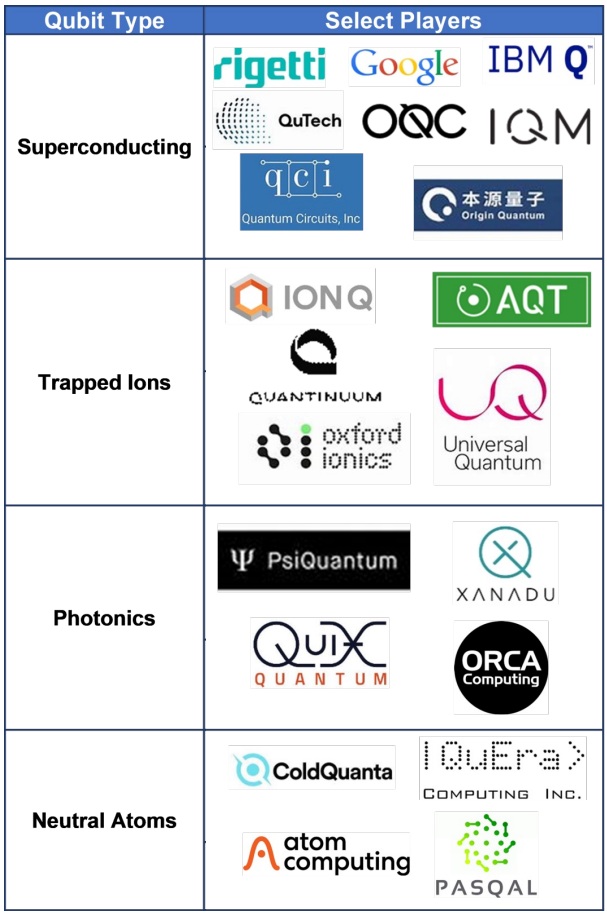


Figure 1.5: Overview of the major players over different quantum computing platforms. Taken from [1]

quantum computers based on this platform.

- **Photonic Qubits:** Using photons as qubits, this platform leverages the properties of light for quantum computation. Photonic systems can operate at room temperature and integrate well with existing optical technologies, together with easy scalability of the number of qubits, these platforms are also gaining popularity in recent years. The main challenge of photonic quantum computing regards the storing of the qubits, as it requires the storing of the photons themselves. Furthermore, photonic quantum computing will likely require measurement-based quantum computing, which requires readdressing of most of the algorithms and many more qubits. Xanadu and PsiQuantum are key companies working on photonic quantum computers, but yet not to be programmable as a general-task quantum computer.

- **Quantum Annealers:** Quantum annealing is a specific type of quantum computation used for optimization problems. It leverages quantum tunneling and superposition to find the global minimum of a given function. D-Wave Systems is a leading company in developing quantum annealers, which are designed to solve specific optimization problems more efficiently than classical computers.

An overview of all these platforms and more can be found in [1], where an extension of Fig. 1.5 is found.

Independently of the platform, current quantum computers are considered to be in the Noisy

Intermediate-Scale Quantum (NISQ) era, which is characterized by a limited number of qubits prone to noise. In this regime noise is a significant challenge, impacting the accuracy and reliability of quantum computations. NISQ quantum computers operate on qubits which are highly sensitive to their environment.

While attempts to attenuate errors are present, in this regime fault-tolerant quantum error correction is not available, thus any *error mitigation*[63, 144] technique, aiming at fractionally reducing the final noise, is going to be cursed by the exponentially increasing error rate in terms of the length of the circuit.

In quantum computing, noise can arise from various sources:

- **Decoherence:** This occurs when qubits interact with their environment, losing their quantum coherence. Decoherence times (T1 and T2) measure how long qubits maintain their state before becoming entangled with their surroundings and losing information.
- **Gate Errors:** Imperfect implementation of quantum gates can introduce errors during quantum operations. Even slight deviations in gate operations can significantly affect computation outcomes.
- **Measurement Errors:** Errors can also occur during the readout process, where the quantum state of a qubit is measured. Measurement errors can lead to incorrect interpretation of qubit states.

A full treatment of error mitigation and correction techniques is out of the scope of this work, however, we refer to Ref. [39] for an extensive explanation of the topic.

1.2.3 VQE

Various methods have been developed to solve the electronic structure problem using quantum computers, with quantum phase estimation being the most well-known[83]. However, considering the field's current transition into the recently introduced NISQ (Noisy Intermediate-Scale Quantum) era, this thesis emphasizes the VQE algorithm, which stands out as one of the most promising approaches for today's quantum devices[54, 99, 113].

The VQE is a hybrid quantum-classical algorithm to find variational solutions for eigenvalue or optimization problems. It is a hybrid algorithm in the sense that it splits the computational effort over both classical and quantum resources, with the classical part taking care of the variational update of parameters and the quantum subroutine handling evaluation of a cost function representing the optimization problem.

The algorithm is based on the *Rayleigh-Ritz* variational principle [128], which states that, given an Hamiltonian operator H and its relative ground state energy E_{GS}, any wavefunction $|\psi\rangle \in \mathcal{H}$ satisfies the following inequality:

$$\langle\psi| H |\psi\rangle \geq E_{GS} \tag{1.49}$$

Given the property shown by (1.49), the concept of the VQE algorithm is then to variationally find an optimal quantum circuit implementing $|\psi\rangle$, where the variational property is implemented in the quantum circuit itself. Eventually, the state implemented is without loss of generality representable by (1.50) ,

$$|\psi\rangle = \left|\psi(\vec{\theta})\right\rangle = U(\vec{\theta})\,|\mathbf{0}\rangle \tag{1.50}$$

so that equation (1.49) becomes

$$\langle\mathbf{0}|\,U^{\dagger}(\vec{\theta})HU(\vec{\theta})\,|\mathbf{0}\rangle \geq E_{GS} \tag{1.51}$$

where the $\vec{\theta} \in \mathbb{R}^k$ is a vector of k real parameters and the unitary operator U represents the exact parametrized quantum circuit of the form $U(\vec{\theta}) = U_{N-1}(\theta_{N-1})\ldots U_k(\theta_k)\ldots U_0(\theta_0)$.

Before getting into details of how the circuit is constructed, let us consider the steps defining the VQE algorithm, the main idea is to use eq (1.51) to get an approximation of the groundstate energy

$$\min_{\vec{\theta}\in\mathbb{R}^k} \langle\mathbf{0}|\,U^{\dagger}(\vec{\theta})HU(\vec{\theta})\,|\mathbf{0}\rangle \sim E_{GS} \tag{1.52}$$

which, if the Hamiltonian operator is not degenerate, implies:

$$|\psi_{GS}\rangle = U(\vec{\theta}_{GS})\,|\mathbf{0}\rangle \qquad with \quad \vec{\theta}_{GS} = \operatorname{argmin}_{\vec{\theta}\in\mathbb{R}^k} \langle\mathbf{0}|\,U^{\dagger}(\vec{\theta})HU(\vec{\theta})\,|\mathbf{0}\rangle \tag{1.53}$$

One important aspect of the Hamiltonian operator H is that it must not necessarily be an actual Hamiltonian operator from a real system, as in reality it can be any Hermitian function with compact support on the Hilbert space. This gives the VQE algorithm a structure similar to those used in other areas, such as quantum machine learning [24, 112], where the cost function, in that case, represents the problem that needs to be learned. In general, all these VQE-*like* algorithms are considered together in the more general class of Variational Quantum Algorithms, even if this class comprehends any variational algorithm running on a quantum device regardless of its similarity with the VQE. Given that H must be a Hermitian operator, it can be proven to be always representable in terms of linear composition of tensor products of the Pauli matrices [92].

$$H = \sum_j h_j P_j = \sum_j h_j \bigotimes_i \sigma_i^j$$

This decomposition itself might not be a trivial task, but it is assured to exist. Overall, the VQE algorithm is composed of three main steps:

1. Prepare the parameterized state $\left|\psi(\vec{\theta})\right\rangle$ on a quantum computer, this is also called Ansatz and is the main focus of this work so more details will follow.

2. Measure the expectation value of the Hamiltonian $\left\langle\psi(\vec{\theta})\right| H \left|\psi(\vec{\theta})\right\rangle$. This is done by exploiting the linearity of expectation values:

$$E = \left\langle\psi(\vec{\theta})\right| \sum_j h_j P_j \left|\psi(\vec{\theta})\right\rangle \tag{1.54}$$

   Essentially, the expectation values of the Pauli strings composing the Hamiltonian are measured by repeating the measurement process many times. This step is executed on the quantum computer, and the number of measurements required to determine the energy with precision ϵ scales as $\mathcal{O}(1/\epsilon^2)$. Overall, while this is not an intrinsically difficult task within the

quantum computation, it is often the main bottleneck of a quantum experiment. In quantum chemistry in particular, we will see that the operator H is composed of $O(N^4)$ terms, so that once measurement repetitions are considered the number of quantum circuits needed is orders of magnitude more than what could be feasible [147] with the current execution time of the gates. Although this load can be parallelized over many devices, great focus is laid on the development of methods that decrease as much as possible the need for so many measurements, we cite here the work regarding commuting sets of Pauli operators [41] and the use of informationally-complete Positive Operator Value Matrices (POVM). [57]

3. The energy value is reconstructed from the measured Pauli strings, weighted with their coefficients: $E = \sum_j h_j \langle P_j | P_j \rangle$. The vector of parameters $\vec{\theta}$ is updated via a calculation on a classical computer. This step, involving a classical optimization, is crucial for the quality of the results of a VQE and is based on classical optimization algorithms. Common optimizers are the Broyden–Fletcher–Goldfarb–Shanno (BFGS) [36] algorithm and the constrained optimization by linear approximation (COBYLA) [117].

In summary, before the algorithm begins, the parametrized quantum circuit is predefined to enable the iterative process. During each iteration i, the energy of the quantum circuit, determined by the parameters $\vec{\theta}_i$, is measured and fed back to the classical optimizer. The optimizer then computes and provides updated parameters for the next iteration, $i + 1$.

Circuit Ansatz

As already introduced, the so-called Ansatz for a VQE algorithm is the circuit describing the parametrized evolution from the initial state $|\mathbf{0}\rangle$ to the final state $\left|\psi(\vec{\theta})\right\rangle = U(\vec{\theta})\,|\mathbf{0}\rangle$. The precise composition of the operator U is vital for the quality of the results of the VQE, and while the algorithm itself takes care of the variational optimization of the parameters, up to now there has not been a consensus on a clear strategy on how to define optimal parametrized unitaries, but rather, as the name Ansatz suggests, different kinds of guesses. For quantum chemistry, which is the main focus of this work, the main classification is between hardware-efficient or chemically inspired Ansätze, with some other falling in a hybrid class (see sec.5). Chemically inspired Ansätze involves adapting classical computational chemistry algorithms for quantum computers. The most adopted example is the Unitary Coupled Cluster (UCC) Ansatz, inspired by the Coupled Cluster (CC) method [15, 68]:

$$\left|\psi(\vec{\theta})\right\rangle = e^{T(\vec{\theta}) - T^\dagger(\vec{\theta})}\,|\psi_{ref}\rangle \tag{1.55}$$

where $T(\vec{\theta}) = \sum_i T_i$ includes systematic expansions approximating the exact many-electron function:

$$T_1 = \sum_{i \in unocc} \sum_{j \in occ} \theta_{ij} a_i^\dagger a_j \; T_2 = \sum_{i,l \in unocc} \sum_{j,m \in occ} \theta_{iljm} a_i^\dagger a_l^\dagger a_j a_m + \dots \tag{1.56}$$

where the $a_i^\dagger, a_i$ are the raising and lowering operators mapped on a quantum computer with a technique shown later in section (2.2), *occ* and *unocc* denote occupied and unoccupied states, respectively. The UCC Ansatz is often considered only up to a specific excitation level, usually single and double excitations (UCCSD) [14, 126]. The possible number of double excitations can be shown to scale as roughly as $\mathcal{O}((N-M)^2 M^2)$,[87] with N representing the number of spin-orbitals and M the number of electrons. This polynomial scaling implies that even for simple molecules, this approach might not be feasible in terms of gate count. For this reason, many improvements and

modifications to the UCC Ansatz have been developed to reduce the number of gates and obtain shorter circuits [87, 101, 142, 155]. On the positive side for the UCC Ansatz, we have that given its chemical intuition of the problem, it benefits from good convergence properties.

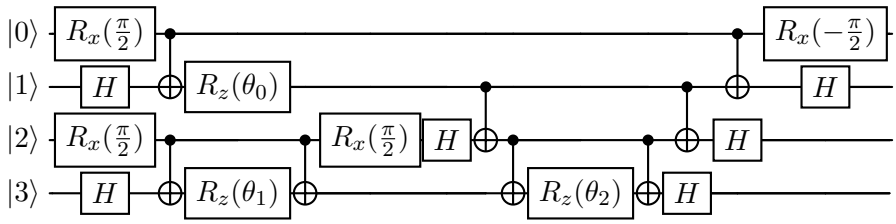


Figure 1.6: Example for the UCCSD Ansatz for an H_2 molecule composed of 4 qubits

The same cannot be said about the second strategy, the hardware-efficient or heuristic Ansätze. Consisting of shallow circuits directly compiled in terms of simple gates as the one available in[76], these Ansätze implement circuits that are similar to other variational quantum algorithms [31] and do not consider any problem-related structure. The task is then to implement circuits that are as simple as possible for the relatively limited quantum devices available currently. A typical heuristic Ansatz consists of repeated blocks with single-qubit rotation gates followed by entangling gates of choice, arranged when possible according to the native hardware connectivity. This same structure is then repeated multiple times to improve the possibility of convergence of the variational algorithm.

As introduced, these Ansätze suffer from bad convergence of the parameters, as the structure as in 1.7 does not allow any guidance, this phenomenon is studied as the phenomenon of the barren plateaux [86] and it is shared between all the applications of this kind of VQA.

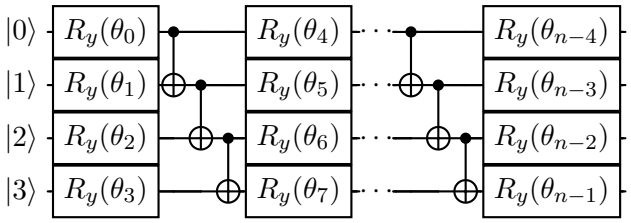


Figure 1.7: Example of a Hardware-efficient Ansatz for a 4-qubit system. Each repeated block is composed of rotations around y-axes on each qubit and a series of CNOTs entangler gates arranged in a ladder.

The part of the present work dedicated to quantum computing algorithms will mainly be focused on the variational quantum eigensolver, with its center being QIDA-*like* Ansätze (5,6). The overall aim will be to find common ground between the two kinds of Ansätze, able to have Hardware Efficient but yet chemically guided circuits.

Chapter 2

Quantum Chemistry

Quantum Chemistry (QCh) is an interdisciplinary field that merges principles of quantum mechanics with chemical systems, aiming to provide a comprehensive understanding of the behavior and interactions of atoms and molecules at the quantum level. The motivation for the development and expansion of QCh is rooted in the need to predict and explain the properties of matter with a high degree of accuracy, surpassing the capabilities of classical theories and entering the realm of quantum physics.

One of the fundamental motivations in QCh is the desire to understand the electronic structure of molecules, which determines a wide range of molecular properties, including their reactivity, stability, and spectra. Classical approaches, while useful, often fail to account for the complex interactions between electrons, particularly in larger or more complex systems. QCh addresses this by applying the principles of quantum mechanics, which consider the wave-like nature of electrons and their probabilistic behavior.

This motivation is generally justified by the need to develop new materials and drugs, as QCh provides insights into molecular interactions at an atomic level, enabling the rational design of new compounds with desired properties. This reduces the need for experimental realization of the compound, and overall accelerates the drug development process and reduces costs. On a deeper level, QCh also addresses fundamental scientific questions about the nature of chemical bonds, reaction mechanisms, and the behavior of matter under extreme conditions, which pushes the boundaries of what is known in chemistry and physics.

The problem of QCh revolves around the Schrödinger equation (2.1), a partial differential equation in quantum mechanics describing the evolution of a system with its Hamiltonian.

$$i\hbar\frac{d}{dt}\left|\psi(t)\right\rangle = \hat{H}\left|\psi(t)\right\rangle \tag{2.1}$$

$$\begin{aligned}\hat{H} = &-\sum_i \frac{\hbar}{2m_e}\nabla_i^2 - \sum_I \frac{\hbar}{2M_I}\nabla_I^2 - \sum_{i,I}\frac{e^2}{4\pi\epsilon_0}\frac{Z_I}{\left|\mathbf{r}_i - \check{\mathbf{R}}_I\right|} + \\ &+\frac{1}{2}\sum_{i\neq j}\frac{e^2}{4\pi\epsilon_0}\frac{1}{|\mathbf{r}_i - \mathbf{r}_j|} + \frac{1}{2}\sum_{I\neq J}\frac{e^2}{4\pi\epsilon_0}\frac{Z_I Z_J}{|\mathbf{R}_I - \mathbf{R}_J|}\end{aligned} \tag{2.2}$$

The Hamiltonian of a chemical system can be written to Eq. (2.2). The first pieces containing ∇ is the kinetic energy of both the electrons i and nuclei I, following the same notation the other

pieces entail the coulomb potential energy of electrons and nuclei between themselves and each other. Bringing the notation to atomic units it becomes:

$$\hat{H} = -\sum_i \frac{\nabla_i^2}{2} - \sum_I \frac{\nabla_I^2}{2M_I'} - \sum_{i,I} \frac{Z_I}{|\mathbf{r}_i - \mathbf{R}_I|} + \frac{1}{2}\sum_{i\neq j} \frac{1}{|\mathbf{r}_i - \mathbf{r}_j|} + \frac{1}{2}\sum_{I\neq J} \frac{Z_I Z_J}{|\mathbf{R}_I - \mathbf{R}_J|} \tag{2.3}$$

This Hamiltonian is clearly time-independent, thus Eq. (2.1) can be rewritten (in its eigenvalue formulation) as:

$$\hat{H}\,|\psi\rangle = E\,|\psi\rangle \tag{2.4}$$

With E being any eigenvalue.

The main focus of QCh is in general to retrieve the lowest energy eigenvalues, so a common task is to retrieve the $E_{GS}, |\psi_{GS}\rangle$ such that:

$$|\psi_{GS}\rangle = \text{argmin}_{\psi\in\mathcal{P}_T} \langle\psi|\,\hat{H}\,|\psi\rangle \qquad with \quad E_{GS} = \langle\psi_{GS}|\,\hat{H}\,|\psi_{GS}\rangle \tag{2.5}$$

This is noticeably the same structure shown in sec. 1.2.3 where the minimization is performed iteratively on a quantum computer, we then begin to see why a quantum device can be useful for QCh. The space of wavefunction where the Hamiltonian is defined is however not yet completely clear, and while it can be brought to the space of qubits, many other passages are necessary to do that. We begin this transition by specifying the starting point. In this formulation, to the ket $|\psi\rangle$ it corresponds a wavefunction $\psi(\mathbf{r},\mathbf{R}) = \langle\mathbf{r},\mathbf{R}|\psi\rangle$ where $\mathbf{r}$ and $\mathbf{R}$ refers respectively to electronic and the atomic positions. The minimization in Eq.(2.5) with this formulation happens over the space of wavefunctions

$$\begin{aligned} &\mathcal{P}_T : \{\Psi \in \mathbb{R}^{3n_e}\text{x}\mathbb{R}^{3n_N} \to \mathbb{C} \subset C^2(\mathbb{R}^{3n_e}\text{x}\mathbb{R}^{3n_N}, \mathbb{C})\} \\ &\text{satisfying :} \\ &\qquad \int^{\mathbb{R}^{3n_e}\text{x}\mathbb{R}^{3n_N}} |\Psi(x)|^2\, d(\mathbb{R}^{3n_e}\text{x}\mathbb{R}^{3n_N}) = 1. \end{aligned} \tag{2.6}$$

Where n_e is the number of electrons, n_N is the number of nuclei and the subscript T refers to the total space.

This formulation in terms of real-space coordinates is quite complex, as each physical body contributes with three coordinates, furthermore, the wavefunction will also have to embody many symmetric properties of the physical system, such as antisymmetry of the wavefunction for electrons exchange and Pauli's exclusion principle. For this reason, it is essential to get rid of any potential complexity that is deemed unnecessary for the resolution of Eq. (2.4). For example, a complete description of the Hamiltonian containing the self-interaction of the nuclei is often not needed for the retrieval of satisfactory E_{GS}. This is expected since the mass of the electrons and of the nuclei are orders of magnitude different. To consider that, we first rewrite

$$\psi(\mathbf{r},\mathbf{R}) = \psi_N(\mathbf{R})\psi_e(\mathbf{r},\mathbf{R}) \tag{2.7}$$

where we split the contributions to the wavefunction into a piece enticing only the nuclei and another contribution for the electrons having a parametrized dependency on the nuclei.

Now, the so-called *Born-Oppenheimer* [32] approximation proposes to ignore the resolution of

Eq. (2.4) for the position of the nuclei and to take their distance only as an external parameter or as part of an iterative procedure. Mathematically this is equivalent to assuming that $|\nabla_{\mathbf{R}}\psi_e(\mathbf{r},\mathbf{R})| \ll |\nabla_{\mathbf{R}}\psi_N(\mathbf{R})|$, the considered Hamiltonian then becomes:

$$\begin{aligned} \hat{H}_e &= -\sum_i \frac{\nabla_i^2}{2} - \sum_{i,I} \frac{Z_I}{|\mathbf{r}_i - \mathbf{R}_I|} + \frac{1}{2}\sum_{i\neq j} \frac{1}{|\mathbf{r}_i - \mathbf{r}_j|} \\ \hat{H} &= \hat{H}_e + C(\mathbf{R}) \end{aligned} \tag{2.8}$$

Where $C(\mathbf{R})$ is now a constant depending on $\mathbf{R}$. For completeness, we also report the time-independent Schrödinger equation (2.4) of this approximated Hamiltonian together with its relative variational principle (2.5).

$$\hat{H}_e |\psi\rangle = E_e |\psi\rangle \tag{2.9}$$

$$|\psi_{GS,e}\rangle = |\psi_{GS,e}\rangle (\mathbf{r},\mathbf{R}) = \text{argmin}_{\psi\in\mathcal{P}_{1,e}} \langle\psi| \hat{H}_e |\psi\rangle \qquad with \quad E_{GS} = \langle\psi_{GS,e}| \hat{H}_e |\psi_{GS,e}\rangle + C(\mathbf{R}) \tag{2.10}$$

$$\begin{aligned} &\mathcal{P}_{T,n_e} : \{\Psi \in \mathbb{R}^{3n_e} \to \mathbb{C} \subset C^2(\mathbb{R}^{3n_e}, \mathbb{C})\} \\ &\text{satisfying :} \\ &\qquad \int^{\mathbb{R}^{3n_e}} |\Psi(x)|^2 \, d(\mathbb{R}^{3n_e}) = 1. \end{aligned} \tag{2.11}$$

This approximation is generally quite reliable, as often the ground state of a molecule does not have superposition over the position of the atoms, thus it is separable by Eq. (1.8), justifying Eq. (2.7). We notice however that this is mainly valid for ground state analysis, as it is more common to see a failing of the approximation for excited states. Since during this work we only work with electronic ground states, it will be considered as a permanent approximation in all our analysis. We can then drop all the different notations introduced in this section to lighten it as much as possible in light of future development, thus for example from now on $E_{GS} = \langle\psi_{GS,e}| \hat{H}_e |\psi_{GS,e}\rangle + C(\mathbf{R}) \Rightarrow E_{GS} = \langle\psi_{GS}| \hat{H} |\psi_{GS}\rangle + C(\mathbf{R})$, retaining the constant $C(\mathbf{R})$ in the definition of the ground state energy. In other words, there won't be any differentiation between the electronic system and the entire one, as we will always mean the former.

2.1 Second Quantization

As we stated already, the wavefunction for a system of multiple electrons must respect the Pauli exclusion principle, which states that no two electrons can occupy the same quantum state simultaneously. This requirement implies that the wavefunction must be antisymmetric under the exchange of any two electrons. However, enforcing this on a wavefunction $\psi \in \mathcal{P}_{T,n_e}$ for the whole electronic system without any helping structure of the space would be rather difficult.

For this reason, a wavefunction for a n_e electrons system is more easily constructed starting

from single-electron (or particle) wavefunction and then combining them in a way so to retain the wanted physical symmetries.

Spin Orbitals and Spatial Orbitals

To describe the state of a single electron, we use a single-electron wavefunction called spin-orbital, which incorporates both spatial and spin components of the electron's state. A spin-orbital $\phi(x)$ is defined as:

$$\phi(x) = \phi(\mathbf{r}, \sigma) \in \mathcal{P} : \{\Psi \in \mathbb{R}^3 \to \mathbb{C} \subset C^2(\mathbb{R}^3, \mathbb{C})\} \mathrm{x} \{\alpha(\sigma), \beta(\sigma)\} \tag{2.12}$$

where $\mathbf{r}$ represents the spatial coordinates and σ represents the spin coordinates.

A spatial orbital $\varphi(\mathbf{r})$ describes the spatial distribution of an electron, and the spin part is described by spin functions $\alpha(\sigma)$ for spin-up and $\beta(\sigma)$ for spin-down. Thus, a spin-orbital can be written as:

$$\phi_i(x) = \varphi_i(\mathbf{r})\alpha(\sigma) \quad \text{or} \quad \phi_i(x) = \varphi_i(\mathbf{r})\beta(\sigma) \tag{2.13}$$

where $\varphi_i(\mathbf{r})$ is the i-th spatial orbital also called molecular orbital, and $\alpha(\sigma)$ or $\beta(\sigma)$ are the spin functions, for more information regarding this division we refer to the complete treating in [11, 128].

Constructing the Slater Determinants (SD)

To construct a wavefunction satisfying the antisymmetry requirement for a system of n_e electrons, we use the SD. The SD is an antisymmetrized product of spin orbitals and ensures that the wavefunction changes sign when any two electrons are exchanged.

For n_e electrons occupying n_e spin orbitals $\phi_1, \phi_2, \ldots, \phi_{n_e}$, the SD is given by:

$$\psi(x_1, x_2, \ldots, x_{n_e}) = \frac{1}{\sqrt{n_e!}} \begin{vmatrix} \phi_1(x_1) & \phi_2(x_1) & \cdots & \phi_{n_e}(x_1) \\ \phi_1(x_2) & \phi_2(x_2) & \cdots & \phi_{n_e}(x_2) \\ \vdots & \vdots & \ddots & \vdots \\ \phi_1(x_{n_e}) & \phi_2(x_{n_e}) & \cdots & \phi_{n_e}(x_{n_e}) \end{vmatrix} \tag{2.14}$$

Here, $\phi_i(x_j)$ represents the value of the i-th spin orbital evaluated at the coordinates of the j-th electron.

Properties of the SD

- **Antisymmetry**: The determinant changes sign when any two rows (or columns) are swapped, ensuring the antisymmetry of the wavefunction:

$$\psi(x_1, \ldots, x_i, \ldots, x_j, \ldots, x_{n_e}) = -\psi(x_1, \ldots, x_j, \ldots, x_i, \ldots, x_{n_e}) \tag{2.15}$$

- **Normalization**: The prefactor $\frac{1}{\sqrt{n_e!}}$ ensures that the wavefunction is properly normalized.
- **Exclusion Principle**: If any two electrons occupy the same spin orbital, the determinant vanishes, reflecting the Pauli exclusion principle.

By constructing the wavefunction as a SD, we ensure that the antisymmetry requirement of the Pauli exclusion principle is inherently satisfied, providing a correct and efficient representation of the quantum state of a multi-electron system in quantum chemistry.

In second quantization, the focus shifts from individual particle coordinates to creation and annihilation operators, which add or remove particles from quantum states. This approach simplifies the treatment of many-body systems, especially when the number of particles is not fixed.

Field Operators

The field operators $\hat{\psi}(\mathbf{r})$ and $\hat{\psi}^\dagger(\mathbf{r})$ annihilate and create an electron at position $\mathbf{r}$, respectively. For fermions, these operators satisfy the anti-commutation relations:

$$\{\hat{\psi}(\mathbf{r}), \hat{\psi}^\dagger(\mathbf{r}')\} = \delta(\mathbf{r} - \mathbf{r}') \tag{2.16}$$

$$\{\hat{\psi}(\mathbf{r}), \hat{\psi}(\mathbf{r}')\} = \{\hat{\psi}^\dagger(\mathbf{r}), \hat{\psi}^\dagger(\mathbf{r}')\} = 0 \tag{2.17}$$

Hamiltonian in Second Quantization

In second quantization, the Hamiltonian for a system of electrons is expressed in terms of these field operators. The Hamiltonian becomes:

$$\hat{H} = \int d\mathbf{r}\, \hat{\psi}^\dagger(\mathbf{r}) \left(-\frac{\nabla^2}{2} - \sum_{i,I} \frac{Z_I}{|\mathbf{r}_i - \mathbf{R}_I|}(\mathbf{r}) \right) \hat{\psi}(\mathbf{r}) + \frac{1}{2} \int d\mathbf{r}\, d\mathbf{r}'\, \hat{\psi}^\dagger(\mathbf{r}) \hat{\psi}^\dagger(\mathbf{r}') \frac{1}{|\mathbf{r} - \mathbf{r}'|} \hat{\psi}(\mathbf{r}') \hat{\psi}(\mathbf{r}) \tag{2.18}$$

Basis Functions and Operators

For practicality, the field operators are expanded in a basis set of spin-orbitals $\{\phi_i(x)\}$:

$$\hat{\psi}(x) = \sum_i \phi_i(x) \hat{a}_i \tag{2.19}$$

$$\hat{\psi}^\dagger(x) = \sum_i \phi_i^*(x) \hat{a}_i^\dagger \tag{2.20}$$

Here, $\hat{a}_i$ and $\hat{a}_i^\dagger$ are the annihilation and creation operators for the i-th single-particle state, satisfying:

$$\{\hat{a}_i, \hat{a}_j\} = \{\hat{a}_i^\dagger, \hat{a}_j^\dagger\} = 0 \qquad \{\hat{a}_i, \hat{a}_j^\dagger\} = \delta_{ij} \tag{2.21}$$

The Hamiltonian in this basis is:

$$\hat{H} = \sum_{ij} h_{ij} \hat{a}_i^\dagger \hat{a}_j + \frac{1}{2} \sum_{ijkl} \langle ij|kl\rangle \hat{a}_i^\dagger \hat{a}_j^\dagger \hat{a}_k \hat{a}_l \tag{2.22}$$

where h_{ij} are the one-body integrals and $\langle ij|kl\rangle$[1] are the two-body integrals, defined as:

$$h_{ij} = \int dx\, \phi_i^*(x) \left(-\frac{\nabla^2}{2} + - \sum_{i,I} \frac{Z_I}{|\mathbf{r}_i - \mathbf{R}_I|}(x) \right) \phi_j(x) \tag{2.23}$$

$$\langle ij|kl\rangle = \int dx_1 \int dx_2\, \phi_i^*(x_1) \phi_j^*(x_2) \frac{1}{|\mathbf{r}_1 - \mathbf{r}_2|} \phi_j(x_2) \phi_l(x_1) \tag{2.24}$$

[1] Following the physicists's notation

Given the framework we constructed, we can now give a more precise definition of the working space.

Definition 3. *Let $\mathcal{P}$ be a single-particle (spin orbitals) Hilbert space. The* Fock space *$\mathcal{F}(\mathcal{P})$ associated with $\mathcal{P}$ is defined as the direct sum of tensor product spaces of $\mathcal{P}$:*

$$\mathcal{F}(\mathcal{P}) = \bigoplus_{n=0}^{\infty} \mathcal{P}^{\otimes n}$$

where:

- *$\mathcal{P}^{\otimes n}$ is the n-fold tensor product of $\mathcal{P}$ with itself.*
- *$\mathcal{P}^{\otimes 0} \cong \mathbb{C}$ is the one-dimensional Hilbert space (interpreted as the vacuum state).*
- *In practice in case of limited particles such as in the fermionic systems under analysis, the direct sum stops at $n = n_e$*

Elements of the Fock space can be written as:

$$|\Psi\rangle = \bigoplus_{n=0}^{\infty} |\psi_n\rangle$$

where $|\psi_n\rangle \in \mathcal{P}^{\otimes n}$.

Depending on the type of particles, the tensor products are symmetrized or antisymmetrized. Given the nature of this work, we will focus on the latter case.

In this formalism, the basis elements of the space consist of all the possible binary strings representing different SD:

$$|k_1, k_2, \ldots, k_N\rangle \tag{2.25}$$

where $k_j = 1$ if the j-th spin-orbital is occupied and $k_j = 0$ otherwise. The creation and annihilation operators act on these states as follows:

$$\hat{a}_i|k_1, \ldots, k_i, \ldots, k_N\rangle = (-1)^{\sum_{j<i} k_j} k_i |k_1, \ldots, 0, \ldots, k_N\rangle \tag{2.26}$$

$$\hat{a}_i^\dagger|k_1, \ldots, k_i, \ldots, k_N\rangle = (-1)^{\sum_{j<i} k_j} (1 - k_i) |k_1, \ldots, 1, \ldots, k_N\rangle \tag{2.27}$$

and thus we can also write

$$|k_1, k_2, \ldots, k_{n_e}\rangle := (a_1^\dagger)^{k_1} (a_2^\dagger)^{k_2} \cdots (a_N^\dagger)^{k_N} |\text{vac}\rangle \tag{2.28}$$

Where N is the number of spin orbitals. We then see that thanks to the definition of Fock space, removing or adding a particle to a SD will leave the involved state inside its space of definition:

$$|\psi\rangle \in \mathcal{F}(\mathcal{P}) \rightarrow a_i / a_i^\dagger |\psi\rangle \in \mathcal{F}(\mathcal{P})$$

In this representation, we can also define the number operator $\hat{N}$ as:

$$\hat{N} = \sum_{i=1}^{N} a_i^\dagger a_i \qquad s.t. \qquad \forall \, |\psi\rangle \in \mathcal{P}^{\otimes n}, \quad \langle\psi| \, \hat{N} \, |\psi\rangle = n \tag{2.29}$$

Which values to the number of particles in the state $|\psi\rangle$.

2.1.1 Basis Set

As we introduced, wavefunctions are constructed from antisymmetrized products of single-particle wavefunctions called spin-orbitals and laying in $\mathcal{P}$. In our previous discussion, however, we did not mention that the process clearly needs a starting basis spanning the space. Namely, we need an approximate reduction from the infinite-dimensional space $\mathcal{P}_{T,n_e}$ defined in (2.11) to a limited number of basis elements for $\mathcal{P}$ In non-periodic quantum chemistry, this is obtained by using functions called Atomic Orbitals (AO), resembling the eigenstate of the Hydrogen atom in either mathematical definition or chemical intuition, as this is one of the few systems that can be solved analytically. In periodic systems, the basis functions (called Plane Waves) are of different shapes and intuition, but the aim is similar.

Overall, these basis functions are essential for the accuracy and convergence of wavefunctions, representing the system's physics with the fewest possible functions, aiming towards the ideal infinite basis set limit. The following treating of this section is partially based on the book by *Surjan* [140], with the mathematical framework mostly based on functional analysis [82].

In general, bases are defined over a discrete spatial object such as atoms, and each of these atoms contributes to the complete single-particle basis. In this framework, the total single-particle space $\mathcal{P}$ is considered as the space by all the basis functions identifying its subsystems:

$$\mathcal{P} = \bigoplus_i \mathcal{P}_i$$

with the sum running over all the subsystems.

Now, each of the separate bases is defined over the same starting spatial space of continuous bounded functions of Eq. (2.11). More precisely, each basis $\mathcal{P}_i$ represents a different incomplete and approximate Hilbert basis of the general space $\mathcal{P}_{T,N_e}$, the direct sum $\bigoplus$ thus defines a still approximate but refined space that spans the union of all the subspaces $\mathcal{P}_i$[2]. However, in this process, the united bases are not yet bound to be orthonormal, as they are different partial descriptions of the same space $\mathcal{P}_{T,N_e}$. In other words, the discretization brought by the different basis sets must be made coherent for different systems as they analyze the same real space.

This detail means that any use of the joint basis $\mathcal{P}_{T,N_e}$ will have to undergo some orthonormalization for its general basis, which is a step included in any computational algorithm based on second quantization. Given that the used basis is always $\mathcal{P}_{T,N_e}$, from now on there will be no need to distinguish between $\mathcal{P}_{T,N_e}$ and $\mathcal{P}$, so that we can drop the subscript for future references.

Let us now review the main type of basis sets used in literature:

1. **Slater-Type Orbitals (STOs)**: These orbitals mimic the hydrogen AO, implementing an exponential decay. The value of the exponential factor ζ is critical. A single-zeta repre-

[2]Technically, each $\mathcal{P}_i$ is a Hilbert space of a given infinite dimensional Hilbertian space which describes the whole real space[3]

sentation employs one ζ value per orbital, whereas an n-zeta representation uses different ζ values for n basis functions per orbital, increasing the radial flexibility at the cost of higher computational complexity.

2. **Gaussian Functions**: Derived from the potential of the three-dimensional harmonic oscillator, these functions facilitate the calculation of two-electron integrals. Gaussian-type orbitals (GTOs) are more localized than STOs due to their quadratic dependence on the radial distance r, requiring multiple functions for accurate representation but allowing efficient evaluation of molecular integrals.

Practical Implementations of Basis Sets
Among the simplest basis sets used are the STO-nG sets, where each atomic orbital is approximated by n Gaussian functions, blending STO properties with the computational advantages of GTOs. Split-Valence sets, like the 6-31G, provide greater radial flexibility for valence orbitals; core orbitals are described by an STO formed from a combination of 6 GTOs, while valence orbitals use a double-zeta approach, with more localized orbitals formed from 3 GTOs and more diffuse ones from a single GTO. As for more advanced basis sets that are going to be used later:

- **Correlation-Consistent Basis Sets** [49]. Developed to systematically approach the complete basis set limit, correlation-consistent (CC) basis sets are optimized for efficient handling of electron correlation effects. These sets, identified by the prefix *cc-pV*, expand predictably to improve calculation accuracy progressively, making them ideal for high-level quantum chemical computations that require detailed correlation understanding.

- **Augmented Basis Sets** [79]. The *aug* label in basis sets typically indicates the inclusion of diffuse functions to standard basis sets. These additional functions are crucial for accurately depicting the electron distributions in molecules with loosely bound electrons, such as anions or molecules with significant lone pairs. For example, the basis set named *aug-cc-pVDZ* points to an augmentation of the correlation-consistent polarized valence double zeta basis set with diffuse functions, enhancing the description of electron delocalization.

- **Def2 Basis Sets** [160]. Primarily developed for DFT applications, these tools strike a crucial balance between speed and accuracy in modern quantum chemistry. They are specially designed with features like polarization and diffuse functions, which are essential for accurately representing the electronic environment of molecules. This design ensures that they effectively manage the complex electron interactions often seen in DFT. The def2 series offers a range of detail levels, from double to quadruple zeta, allowing for an adjustment of the computational effort according to the needs, and enabling efficient computation for a wide variety of molecular sizes and complexities.

2.1.2 Molecular Orbitals choice

While ideally any basis $\mathcal{P}$ which satisfies the mathematical properties listed in previous sections is a valid basis, in practice the performance of most approximate computational methods that aim to solve Eq. (2.4)might depend on the actual choice of basis.

For this reason, we list here three main kinds of orbitals that are used in the literature and in the following of our work. These orbitals are usually combination of the previously listed ones and are constructed tailoring them to the specific chemical system under scope, for this reason are referred to as Molecular Orbitals (MO), while referring to the former ones as Atomic Orbitals (AO).

Hartree-Fock Canonical Orbitals

These orbitals are obtained with a procedure discussed in the following as a result of a variational optimization.

Properties:

- **Energy-Ordered:** They are eigenfunctions of an energy-like operator called Fock operator, thus they can be ordered.
- **Delocalization:** These orbitals are generally delocalized over the entire molecule.

Chemical Intuition:

- **Energy ordering:** The energy ordering helps and maintains the intuition typical of quantum mechanics of defined energy levels. Lower energy orbitals are typically fully occupied, while higher energy orbitals can indicate possible sites for chemical reactions or excitation, this helps to predict reactivity and stability of a system.
- **Delocalized Nature:** Canonical orbitals often spread over multiple atoms, reflecting delocalized electron density. This behavior often limits the extent of chemical intuition for these orbitals in comparison to other chemical theories such as valence bond or Lewis structures analysis

Natural Orbitals

Let us define first the one-body reduced density operator:

$$\rho_{\psi,kl} = \langle\psi| a_k^\dagger a_l |\psi\rangle , \quad k,l \in [1,N] \subset \mathbb{N} \tag{2.30}$$

Now, given a chemical system of interest and its exact ground state $|\psi_0\rangle$ expressed in an undefined basis $\mathcal{P}$, the Natural Orbitals (NO) are defined as the basis for which $\rho_{\psi_0,kl}$ is a diagonal matrix.

Properties:

- **Occupancy Numbers:** By definition, given that $a_i^\dagger a_i$ is the number operator of orbital i, they provide a value for the occupation number, which is a measure of the probability of electron presence in each orbital.
- **Electron Correlation Representation:** They can effectively represent electron correlation, capturing both dynamic and non-dynamic correlation effects. This is emphasized by the idea of the NO as an optimal choice of orbitals for the expression of the wavefunction [93].

Chemical Intuition:

- **Electron Distribution:** The occupation numbers give a clear indication of electron distribution, helping to understand electron correlation in a molecule.
- **Delocalized Nature:** They are also typically delocalized orbitals, in the same manner as the HFCO.

Localized Orbitals

Localized Orbitals (LO) try to align more closely with classical chemical concepts, making it easier to understand and predict molecular properties and reactivity, thus they try to represent the electronic structure of molecules by associating electrons with specific regions in space, such as around individual atoms or within bonds between atoms.

There are several methods to obtain localized orbitals, with some of the most common being:

- **Boys Localization Method:** [61] This method minimizes the spatial extent of orbitals, leading to orbitals that are as localized as possible.
- **Edmiston-Ruedenberg Method:** [50] This approach maximizes the Coulomb repulsion between localized orbitals, which also results in highly localized electron pairs.
- **Pipek-Mezey Method:** [115] This method localizes orbitals by maximizing the sum of orbital variances in atomic populations, which often results in orbitals that are closely associated with specific atoms or bonds.

These methods use different criteria for localization but ultimately aim to provide a more chemically intuitive representation of the orbital basis $\mathcal{P}$.

Properties:

- **Spatial Localization:** Localized orbitals are concentrated around specific atoms or bonds.
- **Non-Orthogonality:** They may not be strictly orthogonal, but the method using these orbitals then needs to account for it.
- **Non-Unique:** Since it is a general concept encompassing different localizing methods, they are not a unique set of orbitals but rather orbitals with a common character.

Chemical Intuition:

- **Bonding and Lone Pairs:** Localized orbitals provide a clearer picture of bonding in terms of electron pairs, aiding in the visualization of molecular structure.
- **Chemical Reactivity:** They highlight regions of electron density that are important for understanding reactive sites, lone pairs, and bond strengths.
- **Hybridization:** Localized orbitals can illustrate the concept of hybridization, showing how atomic orbitals combine to form bonds in molecules.

2.2 Fermionic-to-qubit Mapping

Wanting to bring the fermionic problem on the quantum computer, as preannounced by section 1.2.3, there are a few steps that need to be considered. In a fermionic space, we have seen how the wavefunctions lay in the Fock space $\mathcal{F}$ and how the physical system is strongly constrained by fermionic symmetries such as Eq. (2.21). If we are to bring the problem on the quantum computer, we will have to maintain the same symmetric structure on the qubit space. By their nature, qubits are both distinguishable and bosonic quantum systems, so that not much of the original structure is maintained.

Let us start with a definition for a general fermionic to qubit mapping.

Definition 4. *Given the Fock space $\mathcal{F}(\mathcal{P})$ of the N single-particle wavefunctions defined over $\mathcal{P}$ and a Hilbert space $\mathcal{H}$ composed of tensor product of k qubits as in Eq.*(1.46)*, with $k \geq N$. Then we define as a* ***fermionic to qubit mapping*** *$\mathcal{K}$ any linear map from $\mathcal{F}(\mathcal{P})$ to $\mathcal{H}$,*

$$\mathcal{K} : \mathcal{F}(\mathcal{P}) \to \mathcal{H} \tag{2.31}$$

This map needs to be an homomorphisms, as necessarily the kernel of $\mathcal{K}$ must consist of only the element 0 or else there could be some elements in $\mathcal{F}(\mathcal{P})$ that are not mapped in $\mathcal{H}$. Furthermore, while any mapping with $k = N$ is a valid mapping, and indeed we are going to work with such a requirement later on, this is strictly not necessary, as one can define mappings with qubit spaces having more qubit than it is needed.

Let us now introduce a way of retrieving valid fermionic-to-qubit mappings such that they satisfy the definition. As an important note, the maps that we are going to introduce first and foremost map the excitation operators $a, a^\dagger$ and not a state itself, but as a state can always be composed of excitation operators by Eq. (2.28), this is by no means restrictive. Furthermore, we also acknowledge that this introduction is heavily influenced by *Miller et al.* [103], but other descriptions of fermionic-to-qubit mappings are possible.

We start by defining an alternative set of operators for describing fermionic systems, the **Majorana operators** m_k, defined in terms of creation and annihilation operators:

$$m_{2j} = a_j^\dagger + a_j, \quad m_{2j+1} = i(a_j^\dagger - a_j) \tag{2.32}$$

These operators are unitary, self-adjoint, and satisfy the anti-commutation relations:

$$\{m_i, m_j\} = 2\delta_{ij} \tag{2.33}$$

So the anticommutation relations of Eq. (2.21) simplify to one single expression. Furthermore, if we want to reverse the notation, we have that:

$$a_j^\dagger = \frac{1}{2}(m_{2j} - i m_{2j+1}) \qquad a_j = \frac{1}{2}(m_{2j} + i m_{2j+1}) \tag{2.34}$$

Due to this construction, it is easier to construct fermionic mapping to respect Eq. (2.34), as it reduces to finding $2N$ algebraically independent anticommuting operators on the N qubit register.

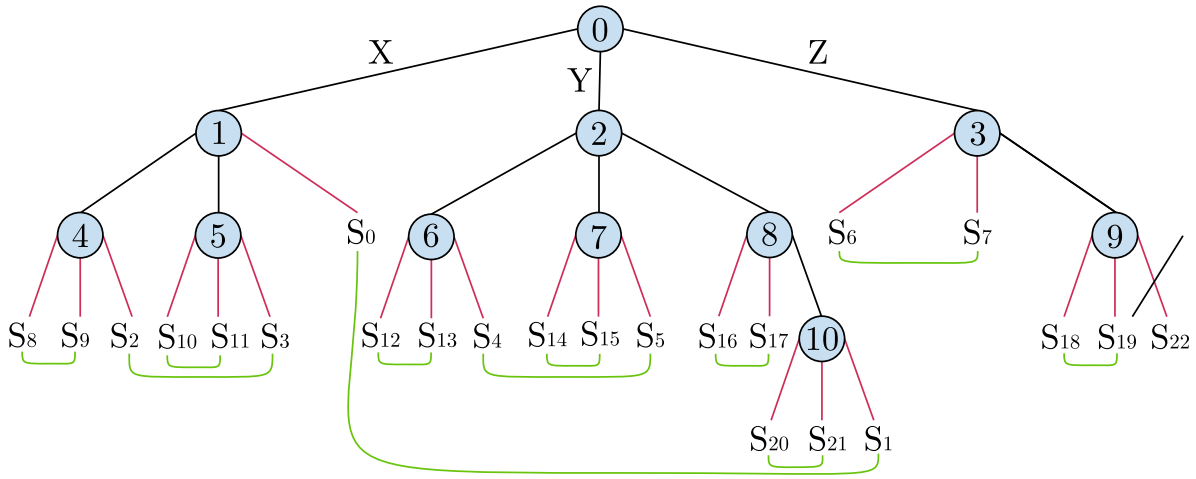


Figure 2.1: Graph of a general ternary tree representing a fermionic-to-qubit mapping of 11 fermionic modes over 11 qubits. Image taken from [103]

For example, in the JW mapping[75] of an N-mode fermionic system, the Majorana operators m_{2j} and m_{2j+1} are mapped to Pauli strings acting on qubits:

$$m_{2j} \mapsto X_j \prod_{k=0}^{j-1} Z_k, \quad m_{2j+1} \mapsto Y_j \prod_{k=0}^{j-1} Z_k \tag{2.35}$$

More generally, and following the construction of [103], a general class of fermion-to-qubit mapping can be represented using Ternary Tree (TT) graphs, a class of tree graph where each node has at most three descendants. In this framework, every node represents a qubit, and every branch from the node is a Pauli operator on that qubit, identity excluded, any possible TT with k nodes represents a mapping over k qubits, fig. 2.1 is an example of what is meant.

Then, given a TT as in fig. 2.1 we can construct the relative mapping by extracting $2K$ Majorana operators from it. This process involves iterating over all the final unconnected branches of the TT, called leaves, and marked by red on the graphical representations. We then call *root* node the only node that is not branched from any other node. For each leaf i, trace the path to the root node, associating for each edge on the path a Pauli operator (X, Y, Z) over the node generating that branch. The resulting Pauli list will constitute the Majorana operator m_i. The so-constructed set of Majorana operators will assure anticommutativity and algebraic independence is achieved by getting rid of one element of this set. The only ambiguity at this point is on the order in which the leaves are analyzed, which regards the pairings of the Majorana operator to then construct the excitation operators as indicated by Eq. (2.34).

By following this procedure, the TT mapping ensures that each fermionic mode is encoded in a specific set of qubits, preserving the product structure and enabling efficient quantum simulation.

2.2.1 Motivation and examples of mappings

While this introduction guarantees the proper construction of a varied class of mapping, it is natural to ask why there is a need for any more research on such a topic, given that mapping as Jordan-Wigner has been formalized almost a century ago in 1928 [75]. The answer is to be looked for in the meaning and properties of each mapping, to do that we are going to list the most famous examples

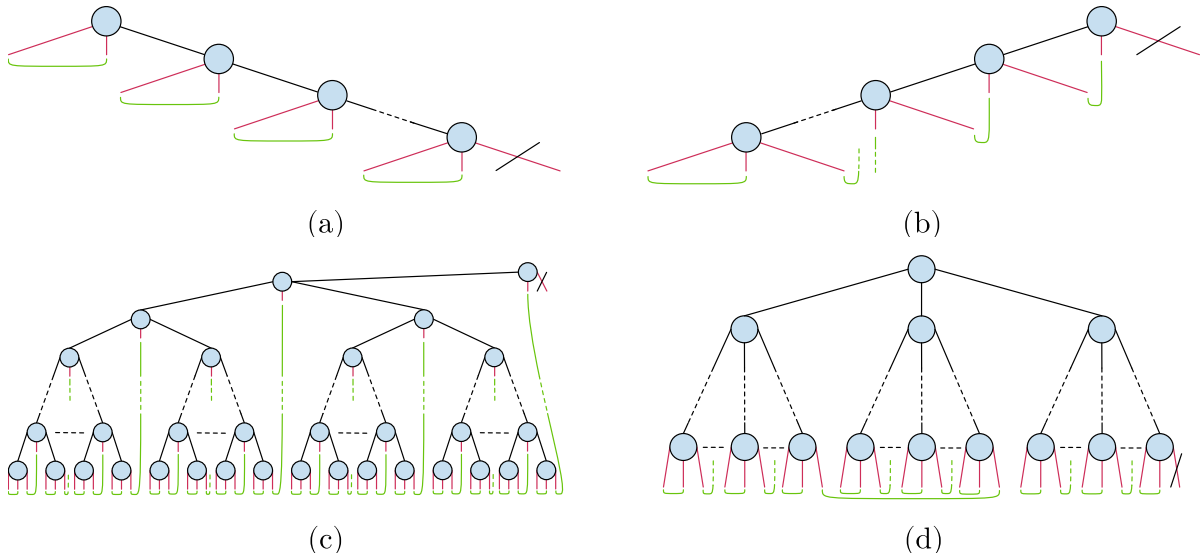


Figure 2.2: Ternary representation of most famous fermionic-to qubit mappings. a) Jordan Wigner Mapping, b) Parity mapping, c) Bravyi-Kitaev, d) JKMN. Image taken from [103]

of mappings and their properties. While the chemical importance of each mapping is determined by how it encodes the fermionic information, the perspective of quantum computing implementations must pass through the definition of Hamming weight of a Pauli string. Take the *Hamming distance* of two Pauli strings

$$\mathcal{D}(\gamma_1, \gamma_2) = \sum_{m=1}^{N} \delta_{\sigma_{1,m},\sigma_{2,m}} \tag{2.36}$$

we call *Hamming weight* of a Pauli string γ the following value:

$$\mathcal{W}(\gamma) = \mathcal{D}(\mathbb{1}^{\otimes N}, \gamma) = \sum_{m=1}^{N} \delta_{\mathbb{1},\sigma_{2,m}} \tag{2.37}$$

The hamming weight is important in quantum computing as for any operator defined in terms of Pauli strings, having a high Hamming weight means having to operate on a higher number of qubits, either through measurement or direct unitary evolution on the qubits involved. We will take the example of the second quantized electronic Hamiltonian of Eq. (2.22) for the following comparison.

Jordan-Wigner

While this mapping[75] certainly satisfies the requirement we introduced at the beginning of the section, it has both drawbacks and strengths. For starting, it is undoubtedly an important mapping in quantum chemistry, as its construction guarantees that the value of encoded by each qubit is the occupation of an assigned spin-orbital in the relative fermionic system.

This means that any SD represented on the quantum device will be the same binary string with the same amplitude in absolute value but with maybe a different phase. This feature will be very useful in future development of the present work (3). Furthermore, as this mapping retains the meaning of the binary values with respect to the fermionic space, it has also been used from the very beginning in both quantum computing and simulation frameworks such as DMRG for quantum chemistry [136, 137].

For the context of this work, since this mapping maintains the meaning of *occupation* over the qubits, it has been used in the following, coherently with the current literature, to establish a direct link with the occupation of the Fock space and to study correlations of part of the system. This is the context of entropy and mutual information introduced in sec. 1.1. In this context, the working space needs to be a separable space, not a property of the starting Fock space, but rather proper of the qubit space.

Analyzing its Hamming weight, given the construction of the JW mapping as in fig. 2.2-a), we can see how the weight of a Majorana operator in terms of its Pauli description can reach at most linear weight in terms of the number of qubit. The components of the electronic Hamiltonian of (2.22) will then have linear weight themselves, which means that one will possibly have to measure all the qubits at once to measure each component.

Parity

In the parity mapping [34], each of the qubit i represents the bitwise ($\oplus$) of all the occupations up to orbital i, so given the fermionic state $|k_1, \dots, k_i, \dots, k_N\rangle$ and the qubit state $|q_1, \dots, q_i, \dots, q_N\rangle$, we have that $q_i = \oplus_{j=1}^{j=i} k_j$.

This mapping has been widely used in recent quantum computing thanks to the fact that if a constant spin number of particles is assumed, as it is often the case for chemical systems, then two of the N qubits are redundant and can be removed. Called *two-qubit reduction*, this feature has been used extensively given the yet limited number of qubits in use in experiments. However, on the practical side, as we can see from fig. 2.2-b), this mapping also has a linear scaling of average Hamming weight in terms of qubits, and thus it has the same disadvantages of the Jordan Wigner mapping, without the advantage of giving information on the chemical occupations of the orbitals. Ideally, this would not be the case, as the relation between the fermionic state and quantum state explained above is totally invertible, however, in the presence of noise this inversion would be quite a challenge, and thus not really implementable. Furthermore, we will see that typical wavefunctions for quantum chemistry involve a sparse type of correlation on the quantum device in a Jordan-Wigner mapping, this is no longer the case with the Parity mapping, since the occupation of a single qubit depends on the occupation of all the previous one.

Bravyi-Kitaev

This mapping [35] implements a base-2 logarithmic Hamming weight of the Pauli strings in terms of the qubit, as represented by fig. 2.2. Introduced before the arrival of TT representation, it has been one of the most advanced mappings for the last two decades.

While it implements a favorable scaling, it suffers from a loss of chemical meaning of the states and a sparse correlation display for quantum chemistry's wavefunctions, yet not as badly as the parity mapping.

JKMN

Introduced in the context of TT [74], it can be proved to be the optimal mapping as far as Hamming weight is concerned. It scales with a base-3 logarithm, and it loses even more both the chemical meaning of a single state and the correlation for quantum chemistry's wavefunctions.

Future prospects For the context of fermionic-to-qubit mapping, much work is still possible and there is still room for advancement. For example, one can implement optimized mappings that concentrate the correlations over fewer qubits as possible [111]. On the other hand, there are also

attempts to use mappings with $k > N$, where the number of qubits is greater than the fermionic modes.

Other works tried to pose the question relative to the mapping only of a determined set of achievable states so as to encode only the states really needed on the quantum device, such as for example a conserved number of particles [65]. In particular, this last attempt revealed that such an operation might involve non-Clifford operations as the one introduced in 1.1, as it is there shown how the passage from the Jordan-Wigner mapping to the Parity mapping only uses the Clifford CNOT gates to implement the bitwise operation. It is then natural to ask whether **all** the mappings representable by a TT can be linked by a Clifford unitary, and to the best of the author's knowledge it is seemingly the current consensus in the field.

Other possible developments pose to question the actual importance of the Hamming weight as the only metric for the optimality of a mapping. As we briefly discussed, any fermionic state will have a different representation depending on the mapping and as such a different correlation pattern. While this property doesn't imply major theoretical differences, it means that different noise models will impact the performance of the quantum algorithm approaching the targeted state differently. Studies on this matter are planned but unfortunately only at an early stage of implementation.

2.3 Classical computational method

We have seen with Eq. (2.9) how we can rewrite the electronic schrödinger's equation to a time-independent eigenvalue problem which solves only for the electrons. We have also seen the structure of the space over which the solution is expected to lie ($\mathcal{F}(\mathcal{P})$). In this section, we will show how this solution is approached by modern computational methods.

We will start with the Hartree-Fock (HF) [66] method, to then develop on them with perturbative (MP2) and Configuration Interaction (CI) based methods. We will then slightly reiterate over the VQE and DMRG methods introduced in sec.1.2.3 and sec.1.1.1 respectively as to discuss them in the context of what explained in this section.

2.3.1 Hartree Fock

The core idea of the Hartree-Fock (HF) method is to reduce the many-body problem to a set of single-body problems. Doing so, the procedure will find the set of single particle spin-orbitals whose occupation with the N electrons of the system minimizes (2.22). In other words, we are looking at the basis of $\mathcal{P}$ whose first N orbitals give rise to the minimizing Slater determinant.

The HF equation

The method involves solving the HF equations, which are derived with the variational principle applied on the Hamiltonian of (2.22) under the assumption that the target wavefunction ψ is a single Slater determinant as in (2.14) with orbitals ϕ_i.

The HF equations are then obtained by minimizing the energy functional with respect to vari-

ations in the orbitals. The energy functional is given by:

$$\begin{aligned} E[\{\phi_i\}] &= \sum_i \langle \phi_i | \, h \, | \phi_i \rangle + \frac{1}{2} \sum_{i,j} (J_{ij} - K_{ij}) \\ J_{ij} &= \langle \phi_i \phi_j | \, \mathbf{r}_{12} \, | \phi_i \phi_j \rangle \\ K_{ij} &= \langle \phi_i \phi_j | \, \mathbf{r}_{12} \, | \phi_j \phi_i \rangle \end{aligned} \tag{2.38}$$

where h is the one-electron Hamiltonian and $\mathbf{r}_{12}$ is the two-electron repulsion operator, derived by Eq. (2.22), they are defined as:

$$h(\mathbf{r}) = -\frac{1}{2}\nabla^2 - \sum_A \frac{Z_A}{|\mathbf{r} - \mathbf{R}_A|} \qquad \mathbf{r}_{12}(\mathbf{r}_1, \mathbf{r}_2) = \frac{1}{|\mathbf{r}_1 - \mathbf{r}_2|} \tag{2.39}$$

Minimizing this functional leads to the eigenvalue equations, also called HF equations:

$$\hat{f}\phi_i = \epsilon_i \phi_i \tag{2.40}$$

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

$$\hat{f} = \hat{h} + \sum_j (\hat{J}_j - \hat{K}_j) = \hat{h} + \hat{\nu}^{\mathrm{HF}} \tag{2.41}$$

where now $\hat{h}$ is the one-electron Hamiltonian, $\hat{J}_j$ is the Coulomb operator, and $\hat{K}_j$ is the exchange operator.

$$\hat{J}_j \phi_i(\mathbf{r}) = \left(\int \frac{|\phi_j(\mathbf{r}')|^2}{|\mathbf{r} - \mathbf{r}'|} d\mathbf{r}' \right) \phi_i(\mathbf{r}) \tag{2.42}$$

$$\hat{K}_j \phi_i(\mathbf{r}) = \left(\int \frac{\phi_j^*(\mathbf{r}')\phi_i(\mathbf{r}')}{|\mathbf{r} - \mathbf{r}'|} d\mathbf{r}' \right) \phi_j(\mathbf{r}) \tag{2.43}$$

Having found the HF equations, the iterative procedure to find the so called Hartree-Fock Canonical Orbitals (HFCO) is the following:

1. **Initial Guess**: Start with an initial guess for the set of orbitals.
2. **Construct Fock Matrix**: Use the guessed orbitals to construct the Fock matrix.
3. **Solve Eigenvalue Problem**: Diagonalize the Fock matrix to obtain a new set of orbitals and orbital energies.
4. **Check Convergence**: Compare the new orbitals with the previous set. If they are sufficiently similar, the solution has converged. If not, repeat steps 2-4 using the new orbitals.

This iterative process usually continues until a chosen convergence threshold in energy is reached.

The HF method has allowed in the past the first major breakthrough in computational QCh, which until then was still limited by a mainly analytical approach. Nevertheless, the single determinant description of a wavefunction has its limitations, and it is not able to approach indefinitely the real ground state of Eq. (2.22). The minimal limit of the energy retrieved by a single determinant approach for a given system at a complete basis limit is called HF-limit. With a fixed basis

set instead, the missing energy that one can retrieve by improving over the HF method is called *Correlation energy*, with the name due to the supposed lack of correlation of the HF approximation.

$$E_c = E_{exact} - E_{HF} \tag{2.44}$$

In this definition, one is supposed to know the exact energy E_{exact} of the system, either by experimental means or by exact solution of the Hamiltonian with a method introduced in 2.3.3. The correlation energy is used in literature mainly as a comparison for the performance of general methods, and it will be used in the following to establish the quality of the VQE Ansätze developed in this work. Contrary to what the name suggests, the HF Slater determinant already contains part of the correlation due to antisimmetrization of the wavefunction, however, it is only given by a physical property of *any* fermionic system and enforced by the mathematical structure of the space.

There are two interesting perspectives on HF theory that deserve a mention. While the derivation above is purely analytical and the only approximation is the requirement of a single Slater determinant for the target wavefunction, there is another simpler derivation. It works by analyzing the energy of an electron given the coulomb field of all the other fixed electrons. This is the idea behind the common *mean-field* name of the HF method. Another interesting perspective is obtained by wondering which Hamiltonian, if any, has the HF state as its ground state. It turns out that the sum of Fock operators on n_e electrons is the Hamiltonian we are looking for. The resulting energy is however not a physical energy, but rather the sum of the orbital energies. This last perspective allows first of all to interpret the energy of the HF ground state as a sum of the contribution of different cumulative orbitals filled up in a typical manner, as the chemical intuition would suggest. The other advantage is that it allows to lay the basis for a perturbative treatment of the complete Hamiltonian starting from the partial Fock Hamiltonian.

2.3.2 Möller-Plesset Perturbation Theory

Following this line of thought, we introduce a method to solve perturbatively the missing part of the Hamiltonian.

This method, which is also called Møller-Plesset perturbation theory of second order (MP2) [105] makes extensive use of the Rayleight-Schrödinger Perturbation theory, whose derivation is out of the scope of this work. For this reason, we give the derivation of the equation relative to perturbation theory as assumed, and further reference for this can be found in [11, 128], while the overall discussion can be found in [11].
Perturbation theory regards the perturbative expansion of a solvable Hamiltonian toward a Hamiltonian whose solution is unknown, in our case, we will call the solvable Hamiltonian H_0:

$$\hat{H}_0 = \sum_{i=1}^{n_e} \hat{f}_i \tag{2.45}$$

Where $\hat{f}_i$ is the fock operator acting on the i electron. The property of this Hamiltonian, as we mentioned in the closure of the previous section, is that the HF-state is the ground state of this Hamiltonian. The total Hamiltonian operator is then:

$$\hat{H} = \hat{H}_0 + \hat{\mathcal{V}} \tag{2.46}$$

Since this Hamiltonian is actually the one in Eq. (2.22), $\mathcal{V}$ is defined by subtraction as:

$$\hat{\mathcal{V}} = \sum_{i<j} r_{ij}^{-1} - \hat{V}^{\mathrm{HF}} = \sum_{i<j} r_{ij}^{-1} - \sum_i \hat{\nu}_i^{\mathrm{HF}} \tag{2.47}$$

Where the single-electron $\hat{\nu}^{\mathrm{HF}}$ was defined in Eq. (2.41).

Given this definition, we can rewrite the results of the findings of Mr̈oller and Plesset describing the system up to second order with the following chain of equations, in which for any quantity of the form $x_i^{(n)}$, i will be the targeted excitation level and n the perturbation order:

$$E_0^{(0)} = \langle 0|\, \hat{H}_0\, |0\rangle = \sum_i \epsilon_i \tag{2.48a}$$

$$E_i^{(1)} = \langle i|\, \hat{\mathcal{V}}\, |i\rangle = -\frac{1}{2}\sum_{a,b}\langle ab||ab\rangle \tag{2.48b}$$

$$|\psi_i^{(1)}\rangle = \sum_{n\neq i} \frac{\langle i|\, \hat{\mathcal{V}}\, |n\rangle}{E_i^{(0)} - E_n^{(0)}}\, |n\rangle = \sum_{a<b,r<s} (-1)^{perm(a,b,r,s)} \frac{\langle ab||rs\rangle}{\epsilon_a + \epsilon_b - \epsilon_r - \epsilon_s} a_s^\dagger a_r^\dagger a_b a_a |\mathbf{HF}\rangle \tag{2.48c}$$

$$E_i^{(2)} = \sum_{n\neq i} \frac{|\langle i|\hat{\mathcal{V}}|n\rangle|^2}{E_i^{(0)} - E_n^{(0)}} = \sum_{a<b,r<s} \frac{|\langle ab||rs\rangle|^2}{\epsilon_a + \epsilon_b - \epsilon_r - \epsilon_s} \tag{2.48d}$$

$$|\psi_i^{(2)}\rangle = \sum_{n\neq i} \frac{\langle n|\hat{\mathcal{V}}|\psi_i^{(1)}\rangle - E_i^{(1)}\langle n|\psi_i^{(1)}\rangle}{E_i^{(0)} - E_n^{(0)}}\, |n\rangle \tag{2.48e}$$

In these equations, the term $perm(a,b,r,s)$ serves to account for the reordering of the excitation operators in $a_s^\dagger a_r^\dagger a_b a_a$, as what is wanted is just the Slater determinant for the double excitation, thus this factor is equal to the number of permutations needed to put the indexes a, b, r, s in increasing order.

The use of MP2 in QCh allowed an immediate inclusion of most of the correlation energy, with basically the same implementation cost of HF, as the latter scales as $O(N^4)$ while MP2 has a scaling of $O(N^5)$ not quite due to Eq. (2.48c) but rather to the change of basis from the atomic integrals to molecular integrals, namely, the tensor h_2 of the form $\langle\phi_i\phi_j|\, \mathbf{r}_{12}\, |\phi_k\phi_l\rangle$.

2.3.3 Configuration Interaction

In QCh, the name Configuration Interaction (CI), refers to all the algorithms that use a diagonalization within the space of viable *configurations*, a.k.a. Slater determinants. Following the same concepts there exist many branching methods that differ mainly by the kind of configurations that they consider. First and foremost, we formalize the theory behind this method, we then differentiate all the existing branches.

The main idea of configuration interaction is to work within the Fock space $\mathcal{F}(\mathcal{P})$, with $\mathcal{P}$ consisting of the orbitals HF procedure or other guesses of orbitals. However, this space is practically never taken as a whole, as there are always Slater determinants that are not expected to contribute to the particular final wavefunction. As a leading example, we can consider Slater determinants which are formed with a number of electrons different than what the system is known to contain. We can call the reduced space of used configurations $\mathcal{F}_{red}(\mathcal{P})$. The CI algorithms are then all the

algorithms that try to solve the following proposition:

$$|\psi_0\rangle = argmin_{|\psi\rangle \in \mathcal{F}_{red}(\mathcal{P})} \langle\psi|\hat{H}|\psi\rangle \qquad E_0 = \langle\psi_0|\hat{H}|\psi_0\rangle \tag{2.49}$$

Where $\hat{H}$ has been defined in (2.22).
Clearly, this construction follows the same concept of the VQE algorithm with Eq. (1.53). The difference is however that this minimization is performed by classical computations through a diagonalization procedure. To do that we work on the on the basis of the space of $Basis(\mathcal{F}_{red}(\mathcal{P}))$, which are all the possible slater determinants, so to form the matrix:

$$\langle i|\hat{H}|j\rangle, \qquad \forall i,j \in Basis(\mathcal{F}_{red}(\mathcal{P})) \tag{2.50}$$

Eq. (2.50) is then usually diagonalized with iterative algorithms such as the Davidson algorithm [43] or the Lanczos [85] algorithm. These algorithms overall favour the scaling, as they often don't need to actually save in memory the matrix (2.50) and instead only work by repeated application of the Hamiltonian on a guess wavefunction, e.g. $\hat{H}\,|\psi_i\rangle$ for step i.

This technique, together with common precise structure of the basis, often allow a great reduction in the cost of CI algorithms, as the major bottleneck of these procedure becomes the calculation of $\hat{H}\,|\psi_i\rangle$.

We hereby specify all the algorithms that fall in the class of CI based algorithms, explaining the main features, differences and computational costs for all the different cases. In the following we will always refer to subspaces of $\mathcal{F}_{red}(\mathcal{P})$ to characterize the differences between all the CI methods.

- Full CI (FCI), the space is composed of all the possible physical states of the system and thus it gives the exact energy of the system up to basis precision. As mentioned before, there can be still excluded configurations, as the system is expected to have a given number of electrons, so Slater determinants with more (or less) than that number are kept out of the used space. These method has an exponential scaling in the number of basis function as the possible number of configurations grows in the same way. It is used in literature and in the following as metric of comparison to be used in the definition of correlation energy.

- Complete Active Space (CAS)-CI. In this method, one defines a *core* an *active* and a *virtual space*, the selected determinants for the CI are then all the ones that keep the core space occupied and the virtual space unoccupied. Basically, only excitations within the active space are allowed, hence the name. If the active space is the whole space, then it is equivalent to the FCI. The cost is exponential in the number of orbitals in the active space.

- CAS-SCF. This algorithm is based on the same structure of the CAS-CI method, which is a subroutine of the method. It involves optimization of the molecular orbitals performed by evaluating iteratively the CAS-CI energy and its derivatives over orbital change. The cost is dominated by the CAS-CI step, while the number of parameters for the orbital rotation scales as $O(N^2)$.

- CI with Singles and Doubles excitations (CISD). In this method, the space is composed of all the configurations that can be obtained from the HF Slater determinant $|HF\rangle$ *via* singles and doubles excitations. The available configurations are then all the possible obtainable with the

applications of the operators of this kind: $a_i^\dagger a_j \left|HF\right\rangle$ or $a_i^\dagger a_j^\dagger a_k a_l \left|HF\right\rangle$
The indexes of these operators clearly only need to run over occupied orbitals for the lowering operators and virtual unoccupied orbitals for raising operators. This algorithm scales at most as $O(N^6)$, as the most expensive kind of interaction is given by two double excitations interacting via a two-electron body of the Hamiltonian, this would turn into an expression of the form $\left\langle HF\right| (a_i^\dagger a_j^\dagger a_k a_l)(a_m^\dagger a_n^\dagger a_o a_p)(a_q^\dagger a_r^\dagger a_s a_t) \left|HF\right\rangle$, of these 12 indexes, the first four are a second excitation, the middle four are a two-body interaction, and the last four are the other second excitation. For the whole term not to be equal to 0, the set of raising indexes must be equal to the set of lowering indexes, which leads to a maximal number of 6 indexes to run over, plus all the possible set's permutations.

- CI-X. One can naively extend the same principle to other excitation orders. The resulting energy improves and the computational cost gains a factor N^2 for each order above the double excitation

- Multi-Reference Configuration Interaction (MRCI). The concept of configuration interaction can be extended to include selected excitations starting from determinants other than the HF one. In this context, a reference is then the Slater determinants over which one applies the chosen excitation. In the previous example of CISD, the reference was the HF Slater determinant. This is the context of *Single-reference* methods. In this algorithm instead, one starts with a CAS-CI/SCF. From this calculation, the dominant contributions in the found wavefunction are taken (with a threshold) as the new *references*, From now on, the idea is similar to previous implementations. The analogous to the CISD method for the multi-reference case is called MRCISD, which includes all possible Singles and Doubles excitations of all the possible references. The cost is dominated by the same behavior of the CISD algorithm with a multiplicative quadratic factor of the number of references, say χ, then the cost is $O(\chi^2 N^6)$, in general however only a handful of references are taken. In the following by MRCI we will always mean MRCISD, as it is the only use we make of multireference methods.

Another honorable mention is to the CIPSI [70] method, which iteratively expands the space of selection by perturbative methods.

In literature one often sees the name of *Selected*-CI, this name most times refers to more complicated method of selection of the space, however, it can just as well refer to *all* the methods that work by reducing the space with *any* criterion, thus encompassing most of the methods listed here.

For completeness, we can compare the CI methods with other variational methods, such as DMRG or even the VQE method. In the case of DMRG, the peculiarity is that the space is not reduced *a priori* with some guess for the kind of included configurations, but rather it is the algorithm itself that finds the best χ configurations out of the whole space. While this is remarkably useful, the computational cost of the algorithm doesn't allow the parameter χ to be increased to what is usually tackled by other CI methods. As for the VQE, it is the ansatz that defines which states are reachable by the quantum circuit, with increased difficulty in the fact that some states are more easily reachable than others depending again on the ansatz. A caveat in this comparison is that while they are indeed variational, they are not formulated over the actual Fock space but rather over the mapped space of qubits. In both cases what is minimized is a Hamiltonian which

has been mapped with the tools introduced in Sec. 2.2, even if this mapping is rarely a recognized step of the DMRG algorithm.

In general, while the CI methods are of undoubted importance and are still an active, prolific field of research, their widespread adoption as the predominant method is hindered by their size inconsistency, which does not allow to rely on results over bigger and bigger systems or to compare calculations for different-sized systems.

Part II

Developments

Chapter 3

SparQ

In the context of non-tensor network quantum chemistry, up to the introduction of the present work, there hasn't been a method to perform quantum information theory without writing the complete exponential density matrix. Even precluding from the quantum chemical applications, practical quantum information theory on relevant non-tensor network wavefunctions on more than few dozens of qubits are unachievable. In this chapter we introduce SparQ, a tool designed to efficiently compute fundamental quantum information theory quantities on sparse wavefunctions, whose primary target for now is focused on on post-Hartree-Fock wavefunctions.

3.1 Introduction

In recent years, the rapid growth of Quantum Computing (QC)[157] has highlighted the relevance of quantum information theory tools within computational sciences. [17, 55, 71, 133] This convergence has promoted a reciprocal exchange of ideas between computational sciences and QC, moving us closer to the era of hybrid quantum-classical computing.

In the field of Quantum Chemistry [11, 66], this exchange has enabled for existing algorithms[10, 51, 64, 120, 121] to benefit from the inclusion of chemical concepts, exemplified by the UCCSD ansatz [40, 53, 69, 84, 151] for variational quantum chemistry algorithms. Additionally, advancements in quantum information theory have benefited classical quantum chemistry by enhancing the understanding of chemical properties [12, 27, 47] and introducing tools that focus computational efforts where they are most needed. [136]

Despite increasing interest in quantum information theory analysis within computational sciences, most existing works rely on Tensor Network methods, which facilitate the intuitive handling of wavefunctions to extract quantum information properties. Within this framework, the Density Matrix Renormalization Group (DMRG)[89, 129, 130, 149, 150] algorithm has proven to be a highly capable and versatile tool [20, 29], retrieving high-quality wavefunctions in a description, which can then be analyzed using various techniques to extract quantum information.
However, certain applications of tensor network methods exhibit significant limitations in system size and quality. In quantum chemistry, the non-locality of the electronic Hamiltonian makes DMRG calculations for systems larger than a few hundred qubits impractical, leaving a gap for quantum information theoretical analysis on larger systems.

This work tries to fill this gap by introducing the Sparse Quantum State Analysis (SparQ) tool.

We here propose a method to effectively calculate the most basic tools of quantum information theory for any wavefunction that is sparse in its definition space. The main focus of this procedure is to manipulate quantum chemistry's wavefunctions, coming from Post-Hartree-Fock (HF) methods, directly to the qubit space through a fermion-to-qubit mapping [75, 103]. Furthermore, given the inspiration from which this work arises[96], it falls directly in line with the mentioned concept of bringing QCh and quantum computing closer. From the quantum computing perspective, other tools for sparse simulation of states on the qubit space exist [104, 139], but their purpose is not state analises but rather to simulate a efficient approximate circuit simulators.

The manuscript is organized as follows, at the beginning we will dive deeper into the necessity of a tool different from tensor networks to analyze wavefunctions, justifying it in part with the chemical background that has sparked the idea for this work. Following this track, we will show how to encode fermionic wavefunction to the qubit space.
We then proceed with a detailed explanation of the developed tools, with an analysis of the performance and computational costs of the introduced procedures. We conclude by proposing potential usage of the method and by showing practical applications.

3.2 Wavefunction representations and sparsity in Quantum Chemistry methods

Given the focus of the authors and the method's primary application, this section is dedicated to giving a context to the use of wavefunctions in quantum chemistry and to show how one can map them to the qubit space. However, since this is not essential for understanding the quantum information part of this work, we invite a reader solely interested in analyzing sparse wavefunctions in qubit space to divert over section 3.6 and forward. In quantum chemistry, the MPS-based DMRG[130] has been the only method used for quantum information theoretical analysis. However, the DMRG algorithm cannot handle quantum chemistry applications involving more than a few hundred qubits, which correspond (roughly) to the same amount of spin orbitals. This is due to the scaling of the DMRG algorithm depending heavily on the so-called *bond dimension* χ, with the scaling being lower bounded by $O(N^3\chi^2) + O(N^2\chi^3)$ [154]. This scaling, even if particularly appealing compared to other methods for small to medium sized systems, quickly becomes unmanageable, especially given the growing requirements in terms of χ with growing system sizes.

By contrast, many other Post-HF methods, given their excitation-based structure, achieve a total scaling as low as $O(N^5)$ (for Møller-Plesset perturbation theory of second order (MP2)[105]), which renders possible calculations on up to thousands of spin orbitals ($\sim$qubits). Clearly, such a precision-oriented algorithm as DMRG should not be so unfairly compared to perturbation theory (or other low-quality) methods, however, the possibility of generalizing quantum information analysis by using wavefunctions coming from other methods could extend its reach to previously unfeasible systems. For completeness, we remark that other techniques, based on transition operators, [27, 30, 125] also apply to methods different than DMRG, since the necessary ingredients are only some terms of the n-body operators. However, besides the limited application that they find within other Post-HF methods, a transition-operator-based study can only be undergone for a few qubits, given the complexity of the derivation of these operators.

The stated generalizing purpose raises the point of bringing fermionic wavefunctions defined

on the Fock space ($\mathcal{F}$) to the separable space of the qubits. This conceptually non-trivial task is handled by the next subsection. For what concerns us for the moment, the wavefunctions ($|\psi\rangle$) under scope can always be written as in Eq. (4.1.1).

where each $|\mathbf{i}\rangle$ is a SD, a vector of a basis of $\mathcal{F}$.

By this definition, there could be a number of components in the wavefunction growing exponentially with N, the number of Spin Orbitals (SO). However, most Post-HF methods such as MP2 [105], Configuration Interaction (CI) [11], Coupled Cluster (CC), and their multireference variants [7, 8, 108, 134], employ relatively few states out of the exponentially many in (4.1.1), this is because they usually rely only on a limited number of excitations. For example, in the case of the common double excitations limit, the expansion in the wavefunction of eq (4.1.1) grows as the number of possible double excitations, which scales as $O(n_o^2 n_v^2)$, for n_o representing the occupied orbitals and n_v the unoccupied ones. The actual space is then reduced to a polynomial number of SD when the total space is still of exponential dimension. This discussion is valid for approximate QCh-PHF methods, however even aspiring for better descriptions, one could still argue that sparsity is still a property of the wavefunction in the fermionic space. As an example of this, imposition of the Super Selection Rules (SSR) on the fermionic space such as particles number usually already limits the effective space to a fraction of the exponential. This discussion is valid for approximate QCh-PHF methods, however even aspiring for better descriptions, one could still argue that sparsity is still a property of the wavefunction in the fermionic space. As an example of this, imposition of the Super Selection Rules (SSR) on the fermionic space such as particles number usually already limits the effective space to a fraction of the exponential even in a FCI description. In this context, even QC frameworks can aim to exploit this sparsity, both in simulation[104, 139] and practical run on QPUs[34]. This work relate to the first case, since we handle sparse quantum states in simulation, however, in our case we don't want to simulate circuit but rather just retrieve pratially traced density matrices.

3.2.1 Encoding of fermionic wavefunctions

Beginning with the wavefunction definition in (4.1.1), we demonstrate how to map the corresponding fermionic excitations to Pauli operators in the separable qubits space using a fermionic-to-qubit mapping.

To maintain a consistent notation for both Pauli operators, SD and (4.1.1), $\mathbf{i}$ will represent the binary string with ones indicating occupied fermionic modes. This means that, for a system of N fermionic modes,

$$a_{\mathbf{i}}^\dagger = \prod_{j=1}^{N} \left(a_j^\dagger\right)^{i_j} \qquad i_j = 0, 1 \qquad j = 1, \ldots, N \tag{3.1}$$

Consequently, the operator $a_{\mathbf{i}}^\dagger$ applies the ordered creation operators of the ones in the $\mathbf{i}$ binary string. In this framework, every state belonging to the Fock space can be represented as

$$|\psi\rangle = \sum_{\mathbf{i}=\{0,1\}^N} c_{\mathbf{i}} a_{\mathbf{i}}^\dagger |\varnothing\rangle \tag{3.2}$$

where $|\varnothing\rangle$ is the vacuum state. As an example, for a system described with 4 spin orbitals, the string $\mathbf{i} = \{1010\}$ is associated to the following Slater determinant $a_{\{\mathbf{1010}\}}^\dagger |\varnothing\rangle = a_1^\dagger a_3^\dagger |\varnothing\rangle = |1010\rangle$.

The Fock space generated by N fermionic modes can be mapped to the N-partite Hilbert space(the qubits space) by means of a *fermionic-to-qubit* mapping, referred to as $\mathcal{K}$. The mapping associates to each creation operator $a_j^\dagger$ the corresponding operator in the qubits space $a_{q,j}^\dagger$ for each $j = 1, \ldots, N$. The mapping preserves the Fermi algebra associated to the fermionic modes

$$\begin{aligned} \{a_i, a_j\} &= \{a_{q,i}, a_{q,j}\} \\ \{a_i^\dagger, a_j^\dagger\} &= \{a_{q,i}^\dagger, a_{q,j}^\dagger\} \\ \{a_i, a_j^\dagger\} &= \{a_{q,i}, a_{q,j}^\dagger\} = \delta_{i,j}\mathbb{1} \end{aligned} \tag{3.3}$$

In [103] it is shown how there exists a class of fermionic-to-qubit mappings such that the vacuum state $|\varnothing\rangle$ is mapped to the vector of the computational basis with all zeros $|\mathbf{0}\rangle$. Assuming that the used $\mathcal{K}$ belongs to this class, we summarize the action of the mapping as follows

$$\mathcal{K}(|\varnothing\rangle) = |\mathbf{0}\rangle \tag{3.4}$$

$$\mathcal{K}(a_j^\dagger) = a_{q,j}^\dagger = \frac{1}{2}(\gamma_q^{2j-1} - i\gamma_q^{2j}) \tag{3.5}$$

$$\mathcal{K}(a_j) = a_{q,j} = \frac{1}{2}(\gamma_q^{2j-1} + i\gamma_q^{2j}) \tag{3.6}$$

where γ_q^l are the Majorana strings, i.e. anticommuting operators in the qubit space associated to the Majorana fermions defined in the Fock space. Now, given this framework, we have Majorana operators of the following form:

$$\gamma_q^l = \bigotimes_{m=1}^{N} \sigma_{l,m} \tag{3.7}$$

where $\sigma_{l,m} = X_m, Y_m, Z_m, I_m$ are the Pauli matrices with the identity, acting on qubit $m = 1, \ldots, N$.

To each creation operator we can associate the value of the Hamming distance calculated taking into account the associated Majorana strings Suppose that for each of the N fermionic modes the Hamming distance is greater than one. In this general case, a product of P fermionic modes is required to store 2^P Majorana strings to encode the wavefunction. This is the case of the parity mapping, as can be easily understood by considering the corresponding tree shown in [103]. However, generally, this prohibitive cost can be reduced if just a subset of the maiorana pairs associated to a given excitation operator has Hamming distance greater than 1. More information regarding the procedure for such cases is given in Appendix 3.3.

3.2.2 Jordan-Wigner mapping

The JW mapping[75], one of the most used and ancient mappings, has Hamming distance equal to one for any mode. This fact renders particularly convenient the encoding of a wavefunction in this mapping. Let us first of all show its definition:

$$\sigma_{l,m} = \begin{cases} I_m & \text{if } \lfloor \frac{l}{2} \rfloor > m \\ X_m & \text{if } \lfloor \frac{l+1}{2} \rfloor = m \\ Y_m & \text{if } \lfloor \frac{l}{2} \rfloor = m \\ Z_m & \lfloor \frac{l}{2} \rfloor - m \end{cases} \tag{3.8}$$

Using this definition, Eq. (3.6) assumes the following form:

$$\begin{aligned} a^\dagger_{q,j} &= \bigotimes_{m=1}^{N} \epsilon_{j,m} \quad with \\ \epsilon_{j,m} &= \begin{cases} \sigma_{2j,m}, & \text{if } \sigma_{2j,m} = \sigma_{2j+1,m} \\ \frac{1}{2}(\sigma_{2j,m} - i\sigma_{2j+1,m}), & \text{otherwise} \end{cases} \end{aligned} \tag{3.9}$$

This allows us to rewrite a single SD of Eq. (3.2) as

$$\mathcal{K}(a^\dagger_{\mathbf{i}}) = a^\dagger_{q,\mathbf{i}} = \bigotimes_{m=1}^{N} \prod_{l=1}^{N} (\epsilon_{l,m})^{\mathbf{i}_l} \tag{3.10}$$

We can now rewrite directly Eq. (3.2) as

$$\begin{aligned} |\psi\rangle &= \sum_{\mathbf{i}=\{0,1\}^N} c_{\mathbf{i}} a^\dagger_{q,\mathbf{i}} |\mathbf{0}\rangle = \\ &= \sum_{\mathbf{i}=\{0,1\}^N} c_{\mathbf{i}} \left(\bigotimes_{m=1}^{N} \prod_{l=1}^{N} (\epsilon_{l,m})^{\mathbf{i}_l} \right) |\mathbf{0}\rangle \end{aligned} \tag{3.11}$$

Now, to pave the way for the usage of wavefunctions of this form in the following section 3.6, we write the wavefunction as follow

$$|\psi\rangle = \sum_{\mathbf{i}=\{0,1\}^N} c_{\mathbf{i}} \left(\bigotimes_{m=1}^{N} |\psi\rangle_{\mathbf{i},m} \right) = \sum_{\mathbf{i}=\{0,1\}^N} c_{\mathbf{i}} |\psi\rangle_{\mathbf{i}} \tag{3.12}$$

where

$$|\psi\rangle_{\mathbf{i},m} = \epsilon_{\mathbf{i},m} |0\rangle_m , \quad \epsilon_{\mathbf{i},m} = \prod_{l=1}^{N} (\epsilon_{l,m})^{\mathbf{i}_l} \tag{3.13}$$

$$and \quad |\psi\rangle_{\mathbf{i}} = \bigotimes_{m=1}^{N} |\psi\rangle_{\mathbf{i},m} \tag{3.14}$$

our ket is now written as a sum of tensor products of single-qubit wavefunctions $|\psi\rangle_{i,m}$. The Eq. (3.12) is the final expression for a generic wavefunction represented on a qubit space, where the sum runs over all the possible SD. However, given the focus of the present work on Post-Hartree Fock CI-based methods, this sum only runs over the possible excitations on a reference wavefunction.

One disclaimer, for the simple case of the JW mapping, one could also map the wavefunction by hand, avoiding this lengthy procedure, as the $a^\dagger$ operators for this mapping are quite simple and translate directly Eq. (4.1.1) to a binary expression over the qubit space, so that the following is true:

$$\sum_{\mathbf{i}} c_{\mathbf{i}} \left|\psi\right\rangle_{\mathbf{i}} = \sum_{\mathbf{i}} k_{\mathbf{i}} \left|\mathbf{i}\right\rangle \tag{3.15}$$

where $\left|\mathbf{i}\right\rangle$ is binary ket. On the other hand, this is equivalent to realizing that in (3.9) and (3.13) no superposition is created by any $\epsilon_{l,m} \left|0\right\rangle_m$, as it can be seen by noting that:

$$\epsilon_{\mathbf{i},m} \left|0\right\rangle_m = \begin{cases} \left|0\right\rangle_m & \text{if } \mathbf{i}_m = 0 \\ \pm \left|1\right\rangle_m & \text{if } \mathbf{i}_m = 1 \end{cases} \tag{3.16}$$

So, for each m we never obtain a linear combination of the two states. This will be used in the following to reduce the computational cost of the trace operator.

The long process for mapping the wavefunction must then be considered for more elaborate mappings, which are built upon the technique explained in this section (Appendix 3.3).

3.3 Mapping generalization

Even if the main focus of quantum information theory for quantum chemistry has always been within the JW mapping, one might also be taken in account, for quantum computing reasons [96, 159], different fermion-to qubit mapping. We here generalize the mapping procedure for other fermionic-to-qubit mappings. As we already mentioned in the Majorana operators γ_q^k must anti-commute with each other. Given this anticommuting relation, the Majorana pair must differ in at least one qubit, we can then assert the following:

$$\mathcal{H}(a^\dagger_{q,j}) \geq 1 \tag{3.3.1}$$

where $\mathcal{H}$ is the Hamming distance corresponding to the qubit operator $a^\dagger_{q,j}$ defined above.

Now, while the Jordan-Wigner (JW) mapping, defined as in (3.8) can be easily proven to satisfy equality in (3.3.1) for any pair of Majorana operators, general fermion-to-qubit mappings $\mathcal{K}$ do not have this property. In such cases, instead of contracting and summing together the two single-qubit Pauli operators as we done in Eq. (3.9), we split each pairs of Majorana strings as follow

$$\begin{aligned} &a^\dagger_{q,\mathbf{i}} \left|\mathbf{0}\right\rangle = \prod_{l=1}^{N} (a^\dagger_{q,l})^{\mathbf{i}_l} \left|\mathbf{0}\right\rangle = \\ &(a^\dagger_{q,1})^{\mathbf{i}_1} \dots (a^\dagger_{q,k})^{\mathbf{i}_k} \dots (a^\dagger_{q,N})^{\mathbf{i}_N} \left|\mathbf{0}\right\rangle = \\ &\prod_{l=1}^{N} (\frac{1}{2}\gamma_q^{2l-1} - \frac{i}{2}\gamma_q^{2l})^{\mathbf{i}_l} \left|\mathbf{0}\right\rangle = \\ &\sum_{k=1}^{2^N} d_k \left(\bigotimes_{m=1}^{N} \sigma_{k,m} \right) \left|\mathbf{0}\right\rangle \end{aligned} \tag{3.3.2}$$

where $\mathbf{i}_l = 0, 1$ and $\sigma_{k,m} = X_m, Y_m, Z_m, I_m$, with d_k coefficients depending on the encoding method and fermionic modes appearing in $\mathbf{i}$. Obviously, the summation runs on a number of elements that depends on the Hamming distance associated to each operator, that are no more than 2^N. For practical cases, this number is drastically reduced. For example, for the JW encoding method, this expression reduce to Eq.3.10 because the Hamming distance for each operator is equal to one. For a generic wavefunction

$$|\psi\rangle = \sum_{\mathbf{i}=\{0,1\}^N} c_{\mathbf{i}} a^\dagger_{q,\mathbf{i}} |\mathbf{0}\rangle \tag{3.3.3}$$

we obtain

$$|\psi\rangle = \sum_{\mathbf{i}=\{0,1\}^N} c_{\mathbf{i}} \sum_{k=1}^{2^N} d_k \left(\bigotimes_{m=1}^{N} \sigma_{k,m} \right) |\mathbf{0}\rangle \tag{3.3.4}$$

that become

$$|\psi\rangle = \sum_{j=1}^{4^N} \lambda_j \left(\bigotimes_{m=1}^{N} |\psi\rangle_{j,m} \right) \quad |\psi\rangle_{j,m} = \epsilon_{j,m} |0\rangle_m \tag{3.3.5}$$

where $\epsilon_{j,m}$ are single qubit operators and the λ_j coefficients depends on $c_{\mathbf{i}}$ and d_k ones. We wrote that the summation runs over 4^N terms because this number is an upper bound but, for practical case, this is no more than polynomial: for example, for the JW encoding method, we obtain Eq.3.11, thus we have a number of strings equal to the number of Slater determinants in the wavefunction, with λ_j being the coefficients of the Slater determinants.

On the mapped wavefunction $|\psi\rangle$, obtained from the encoding method $\mathcal{K}$, we execute the calculations explained in section 3.6. It follow that the cost of these operations is proportional to the cost to encode the wavefunction, defined as the cost associated to the most expensive Slater determinant belonging to $|\psi\rangle$. So

$$C_{\mathcal{K}}(|\psi\rangle) = \max_{\mathbf{i} \in |\psi\rangle} 2^{P_{\mathbf{i}}} \tag{3.3.6}$$

where

$$P_{\mathbf{i}} = \sum_{l=1}^{N} \mathbf{i}_l \Theta(\mathcal{H}(a^\dagger_{q,l}) - 1) \tag{3.3.7}$$

with

$$\Theta(x) = \begin{cases} 0 & \text{if x = 0} \\ 1 & \text{if x > 0} \end{cases} \tag{3.3.8}$$

As can been easily seen, this cost is equal to 1 for the JW mapping, as has been introduced above.

3.4 Efficient mapping permutation

Given that fermion-to-qubit mappings are a relatively active field of research [34, 35, 74, 102, 103], we are hopeful that improvement in the understanding of mappings will bring on a further reduction in the cost of encoding wavefunctions for general mappings, either by reducing $\mathcal{C}_{\mathcal{K}}$ or by a complete revision of the method. On this line, we can further show how there are already specific techniques to bring down such costs for specific mapping. It is widely recognized that the Parity mapping serves as the dual to the JW mapping in terms of the information it encodes.

However, it can be shown [65] that at the level of unitary transformation in the qubit space,

the difference between the two can be summarized by a permutation described by the chained application of CNOTs unitary from the first to the last qubit.

$$\begin{aligned} U_{JW\to P} &= CNOT_{N-1,N}\dots CNOT_{i-1,i}\dots CNOT_{0,1} \\ a^\dagger_{P,i} &= U_{JW\to P} a^\dagger_{JW,i} U^\dagger_{JW\to P} \end{aligned} \tag{3.4.1}$$

Considering that each piece in superposition of (3.11) for our qubit wavefunction is generated by $a^\dagger_{q,\mathbf{i}}$, we can rewrite

$$\begin{aligned} a^\dagger_{P,\mathbf{i}} \left|\mathbf{0}\right\rangle_q = U_{JW\to P} a^\dagger_{JW,\mathbf{i}} U^\dagger_{JW\to P} \left|\mathbf{0}\right\rangle = \\ = U_{JW\to P} a^\dagger_{JW,\mathbf{i}} \left|\mathbf{0}\right\rangle = U_{JW\to P}\left(a^\dagger_{JW,\mathbf{i}} \left|\mathbf{0}\right\rangle\right) \end{aligned} \tag{3.4.2}$$

Since the zero state is an eigenstate of the CNOT and thus $U^\dagger_{JW\to P}\left|\mathbf{0}\right\rangle = \left|\mathbf{0}\right\rangle$. At last, since $U_{JW\to P}\left(a^\dagger_{JW,\mathbf{i}}\left|\mathbf{0}\right\rangle\right)$ is a classical bit-wise operation, one can conclude that this operation can be efficiently done in any case at no further multiplicative cost from the JW mapping. Finally, we notice that since the only requirement used to establish (3.4.2) was the pure CNOTs composition of $U_{JW\to P}$, this procedure can be applied to any mapping in the permutation group of the starting mapping, which in this case is the JW mapping, further details can be found in [65].

3.5 Particle-Hole duality

Up to now we encoded the fermionic vacuum state into the computational state with all the qubits in the $|0\rangle$ state, so

$$\left|\varnothing\right\rangle \to \left|\mathbf{0}\right\rangle \tag{3.5.1}$$

To reduce the computational effort needed to encode a wavefunction and perform some operation we take into account the particle-hole transformation. This then brings us to consider the Hartree-Fock state as the vacuum state and to encode it as follow

$$\left|\mathbf{HF}\right\rangle \to \left|\mathbf{0}\right\rangle \tag{3.5.2}$$

It follow that the operators of the fermionic modes $a^\dagger_i$, with $i = 1,\dots,N$, are transformed as follow

$$a^\dagger_i = \begin{cases} a^\dagger_i & \text{if i} > n \\ b_i & \text{if i} \leq n \end{cases} \tag{3.5.3}$$

where b_i are the annihilation operators for hole particles that are defined for $i \leq n$ and $n = \sum_{k=1}^{N} \left\langle \mathbf{HF}\right| a^\dagger_i a_i \left|\mathbf{HF}\right\rangle$. Obviously, these operators obey to the Fermi algebra

$$\begin{aligned} \{b_i, b_j\} = \{a^\dagger_k, b^\dagger_j\} = \{a_k, b^\dagger_j\} = 0 \\ \{b_i, b^\dagger_j\} = \delta_{i,j}\mathbb{1} \end{aligned} \tag{3.5.4}$$

with $i, j = 1, \ldots, n$ and $k = n+1, \ldots, N$. As an example, consider the following state

$$a_l^\dagger a_m^\dagger a_n a_o \left|\mathbf{HF}\right\rangle = a_l^\dagger a_m^\dagger a_n a_o \prod_{i=1}^{n} a_i^\dagger \left|\varnothing\right\rangle \tag{3.5.5}$$

To map it on the qubits space, we should apply $n+4$ operators to the encoded vacuum state $|\mathbf{0}\rangle$. Taking into account the particle-hole transformation, we obtain that

$$a_l^\dagger a_m^\dagger a_n a_o \left|\mathbf{HF}\right\rangle \rightarrow a_{q,l}^\dagger a_{q,m}^\dagger a_{q,n} a_{q,o} \left|\mathbf{0}\right\rangle \tag{3.5.6}$$

so we just apply 4 operators to the vacuum state. The computational cost, in this case, does not depend on the number of particles n in the Hartree-Fock state, substantially reducing the overhead. "‘ This work tries to fill this gap by introducing the Sparse Quantum State Analysis (SparQ) tool. We here propose a method to effectively calculate the most basic tools of quantum information theory for any wavefunction that is sparse in its definition space. The main focus of this procedure is to manipulate quantum chemistry's wavefunctions, coming from Post-Hartree-Fock (HF) methods, directly to the qubit space through a fermion-to-qubit mapping [75, 103]. Furthermore, given the inspiration from which this work arises[96], it falls directly in line with the mentioned concept of bringing QCh and quantum computing closer.

3.6 SparQ

This section describe the procedure to retrieve efficiently quantities such as the trace and mutual information from any sparse wavefunction in qubit space. Hereafter, since we will no longer refer to the Fock space, we will drop the "$_q$" in the equations as we always refer to the qubit space.

For completeness, we pose the attention of the reader to the two ways of representing sparse wavefunctions.

$$|\psi\rangle = \sum_j \lambda_j \left(\bigotimes_{m=1}^{N} |\psi\rangle_{j,m} \right) \tag{3.6.1}$$

The first kind of wavefunction is the one expressed by Eq. (3.6.1), which is composed of a sum of tensor products of single-qubit wavefunctions, these wavefunctions are indexed by j, and they sum to whichever number of components the treated wavefunction has. The second, less general representation is the binary representation (3.6.2), which in previous sections and the following ones has been referred to with the index $\mathbf{i}$, and in which each $|\mathbf{i}\rangle$ is the ket of the relative binary excitation levels of the qubits.

$$|\psi\rangle = \sum_{\mathbf{i}=\{0,1\}^N} k_{\mathbf{i}} \left|\mathbf{i}\right\rangle \tag{3.6.2}$$

In both notations, the main aim is to treat only a limited number of non-zero $\lambda_j / k_{\mathbf{i}}$. Moreover, we notice that at this point the indexing has only counting purposes, as for Eq. (3.6.1) there cannot be any particular order.

To clarify the differences between the two formalisms, we can take as an example the wavefunction on two qubits$|\phi\rangle = \left(\frac{1}{\sqrt{2}}(|0\rangle_0 + |1\rangle_0)\right) \otimes \left(\frac{1}{\sqrt{2}}(|0\rangle_1 + |1\rangle_1)\right)$. Clearly, by calling $|+\rangle_0 =$

$\frac{1}{\sqrt{2}}(|0\rangle_0 + |1\rangle_0)$ we can rewrite $|\phi\rangle = |+\rangle_0 \otimes |+\rangle_1$ following formalism (3.6.1), however, we could just as well expand it in the complete binary form $|\phi\rangle = \frac{1}{2}(|00\rangle + |01\rangle + |10\rangle + |11\rangle)$, which is the formalism of (3.6.2). While it is the same state, it has a different description in the two notations, and also a different computational cost in its expression, both in time and in memory, as the binary form of $|\phi\rangle$ is not sparse anymore in its space of definition. In the case of the example $|\phi\rangle$ one could resort to a local change of computational basis, which can even be done independently for each qubit, however, the division between the two formalisms remains, as this change of basis might not be done independently for each state of Eq. (3.6.1).

Beginning with the first of these two notations, we have shown in appendix 3.3 that any wavefunction can be encoded as Eq. (3.6.1). For fermionic wavefunctions under the Jordan-Wigner mapping, this reduces to equation 3.12, with λ_j becoming the $c_{\mathbf{i}}$ coefficients as in (3.12).

3.6.1 Partial trace and observables

The primary purpose of the procedures explained in this work is to calculate the properties of a quantum state represented as a vector in a separable Hilbert space. The measurable observables include physical quantities that directly depend on the electronic state, such as energy, dipole moment, and magnetization. In this work we mainly focus on the calculation of quantities that are relevant in the field of quantum information, such as the mutual information.

We can write the density matrix of the mapped state as

$$\rho = |\psi\rangle \langle\psi| \tag{3.6.3}$$

that, by using Eq. 3.6.1, becomes

$$\begin{aligned} &|\psi\rangle \langle\psi| = \\ &= \sum_{j,j'} \lambda_j \lambda_{j'}^* \left(\bigotimes_{m=1}^{N} (|\psi\rangle_{j,m} \langle\psi|_{j',m}) \right) \end{aligned} \tag{3.6.4}$$

3.6.2 Partial Trace

We now have the tools required to reduce the dimensionality of the qubit space starting from Eq. (3.6.4). The reduced density matrix corresponding to the qubit k can be obtained as follows:

$$\begin{aligned} &\rho_k = tr(\rho)_{1,\cdots k-1,k+1,\cdots,N} = \sum_{j,j'} \lambda_j \lambda_{j'}^* \\ &tr\left(\bigotimes_{m=1,m\neq k}^{N} (|\psi\rangle_{j,m} \langle\psi|_{j',m}) \right) \cdot |\psi\rangle_{j,k} \langle\psi|_{j',k} \end{aligned} \tag{3.6.5}$$

From this definition, it is clear that the scaling of this operation is quadratic in the number of wavefunction components χ and linear in the number of qubits being traced $N-1$. Generally, the scaling will be $O(\chi^2 N)$

3.6.3 Observables

Measuring observables (O) on the qubit space is a straightforward operation defined by:

$$O = \sum_k o_k \bigotimes_{m=1}^{N} O_{k,m} \tag{3.6.6}$$

Where $O_{k,m}$ is a single qubit operator.

$$\begin{aligned} &\langle\psi| O |\psi\rangle = tr(|\psi\rangle O \langle\psi|) = \\ &= tr\left(\sum_{j,j',k} \lambda_j \lambda^*_{j'} o_k \left(\bigotimes_{m=1}^{N} (|\psi\rangle_{j,m} \langle\psi|_{j',m} O_{k,m})\right)\right) \end{aligned} \tag{3.6.7}$$

We note that this operation has a quadratic cost in the number of states and a linear cost in the number of qubits, as described in Eq. (3.6.5), however, there is now also a multiplicative cost in the number of operators defining O.

3.6.4 Direct tracing

The partial trace equation introduced in (3.6.5) requires evaluating all possible combinations of two states ($\sum_{j,j'}$) in the outer product defining the density matrix (3.6.3), significantly limiting the number of states χ of the wavefunction (3.6.1) that can be considered. Additionally, the traced density matrix produced will have the full dimension of the space, scaling exponentially. This further restricts the maximum dimension of the space that can be processed.

To overcome these limitations, we change the representation for the wavefunction using the binary representation as in Eq. (3.6.2). This approach reduces the cost of tracing from quadratic scaling in χ to linear scaling, significantly lowering the computational cost. We begin by defining the notation, referring back to Eq. (3.6.2), where each of the $|\mathbf{i}\rangle$ represents a ket in a binary computational basis of the qubit space. To compute the partial trace, we must first define how the various subspace divisions are represented within this binary notation.

Given a ket $|\mathbf{i}\rangle$ as in Eq. (3.6.2) and a subset of qubits A with Hilbert space $\mathcal{H}_A$, we represent the reduction of $|\mathbf{i}\rangle$ to the space of the qubits in A as $|\mathbf{i}\rangle_A$, indicating the binary collection of the qubit states with the following:

$$(|\mathbf{i}\rangle)_A = |\mathbf{i}\rangle_A \tag{3.6.8}$$

While working with different subsets of qubits, say A and B, we can still resort to this notation as long as the two subsets of qubits satisfy $A \cap B = \varnothing$. Furthermore, representing with $\mathcal{N}$ the set of all qubits, if $A \cup B = \mathcal{N}$, then $|\mathbf{i}\rangle = |\mathbf{i}\rangle_A \otimes |\mathbf{i}\rangle_B$, or also $|\mathbf{i}\rangle = |\mathbf{i}\rangle_A \otimes |\mathbf{i}\rangle_{\overline{A}}$, since the space A and B are complementary.

With this notation in mind, we can now retake Eq. (1.7) to see how there is only an intrinsic linear cost in the definition, which is brought by the summation over all computational basis states of B.

Furthermore, since we are working with a sparsely defined vector $|\psi\rangle$ (3.6.2) with χ components, the possible reductions to the space B are going to be at most $min(2^{|B|}, \chi)$.

We can now see some differences between this approach and the one in (3.6.4), since in the latter

the state did not have any intrinsic structure nor order, one could only construct the whole density matrix via outer product (3.6.3) and only then work on tracing part of the qubits.

Instead, the aim is now to start with $|\psi\rangle \in \mathcal{H}_{\mathcal{N}}$ and obtain the density matrix ρ_A.

$$\rho_A = tr(|\psi\rangle \langle\psi|)_{\overline{A}} \tag{3.6.9}$$

To obtain a better scaling, and therefore exploiting the linearity of (1.7), the main idea is to consider the density matrix never doing the outer product directly, but working instead on its wavefunction expression whenever the aim is to treat a pure state. Moreover, we note that in Eq. (3.6.5) the resulting traced density matrix could not always be expressed directly as it is in terms of the computational basis in A, or else it would take full exponential form. In this notation instead, we will obtain the density matrix directly in a sparse form linked to the computational basis of A. This will allow any other kind of sparse operation on the density matrix.

Elucidating the pseudo-code scheme reported in Algorithm 1, the procedure consists of the following steps:

Algorithm 1: Line of work for the direct tracing method.

```
1  Direct Trace (|ψ⟩ , A);
   Input  : |ψ⟩ , A
   Output: ρ_A, H_A
2  γ_A ← {};
3  for |i⟩ ∈ |ψ⟩ do
4  |   if |i⟩_A ∉ Keys(γ_A) then
5  |   |   γ_A [|i⟩_A] ← {};
6  |   end
7  |   γ_A [|i⟩_A] ← γ_A [|i⟩_A] ∪ |i⟩_Ā ;        // Insertion of outer-space vector in the set
8  end
9  H_A ← Keys(γ_A) ;                              // Effective Hilbert space of A
10 n_ρA ← #Keys(γ_A);                                          // ← Dim(H_A)
11 allocate ρ_A(n_ρA, n_ρA);                                   // Initialized to zero
12 c_l ← 0;   c_r ← 0;
13 for |a⟩ ∈ H_A do
14 |   c_r ← 0;
15 |   for |a'⟩ ∈ H_A do
16 |   |   for |ā⟩ ∈ γ_A[|a'⟩] do
17 |   |   |   if |ā⟩ ∈ γ_A[|a⟩] then
18 |   |   |   |   |i⟩ ← |a⟩ ⊗ |ā⟩ ; |j⟩ ← |a'⟩ ⊗ |ā⟩;
19 |   |   |   |   ρ_A(c_l, c_r) ← ρ_A(c_l, c_r) + k_i k*_j;        // with k_i as in (3.6.2)
20 |   |   |   end
21 |   |   end
22 |   |   c_r ← c_r + 1;
23 |   end
24 |   c_l ← c_l + 1;
25 end
26 return ρ_A, H_A
```

1. We start by doing an iteration over all the binary strings in $|\psi\rangle$ (3.6.2)(line 3), during the

iterations we save a dictionary γ_A whose keys are all the existent binary strings in $|\mathbf{i}\rangle_A$, the items of the dictionary are a set of $|\mathbf{i}\rangle_{\overline{A}}$.

The dictionary only saves each binary occurrence of A once (line 5), by adding another element to the $\gamma_A[|\mathbf{i}\rangle_A]$ each time there is the same occurrence of a $|\mathbf{i}\rangle_A$(line 7). The binary strings in A, as keys of the dictionary, will then represent the effective space $\mathcal{H}_A$ (line 9).

2. After having initialized to zero the density matrix $\rho_{\mathcal{H}_A}$ to be filled, remembering Eq. (1.7), we find the value of coefficient of the density matrix relative to the $|\mathbf{a}\rangle\langle\mathbf{a}'|$, with $|\mathbf{a}\rangle, |\mathbf{a}'\rangle \in \mathcal{H}_A$, by considering all the extensions $|\overline{\mathbf{a}}\rangle$ to the traced space $\overline{A}$ in common between $\gamma_A[|\mathbf{a}\rangle_A]$ and $\gamma_A[|\mathbf{a}'\rangle_A]$, i.e.:

$$\sum_{|\overline{\mathbf{a}}\rangle \in \gamma_A[|\mathbf{a}\rangle_A] \cup \gamma_A[|\mathbf{a}'\rangle_A]} \langle\overline{a}|\mathbf{i}\rangle \langle\mathbf{j}|\overline{a}\rangle \, k_{\mathbf{i}} k^*_{\mathbf{j}}$$

With $|\mathbf{i}\rangle \leftarrow |\mathbf{a}\rangle \otimes |\overline{\mathbf{a}}\rangle$; $|\mathbf{j}\rangle \leftarrow |\mathbf{a}'\rangle \otimes |\overline{\mathbf{a}}\rangle$ (line $11-25$)

Computational Cost

As shown in with the pseudo-code 1, this implementation of the partial trace involves one cycle over all the states in the superposition (line 3) and three for-loops to define precisely ρ_A. However, iterating over all the $|\mathbf{a}'\rangle$ (which are in total n_{ρ_A}) and$|\overline{\mathbf{a}}\rangle$ entails running over all the existing combinations of $|\mathbf{a}'\rangle \otimes |\overline{\mathbf{a}}\rangle$, which are χ. Therefore, the final time scaling will be $O(\chi n_{\rho_A})$. This scaling does not account for the tools needed to identify, save, and retrieve efficiently the correct subspaces at each step, with the most significant being the call to $\gamma_A[|\mathbf{a}'\rangle]$ and checking the common $|\overline{\mathbf{a}}\rangle$ in lines $16, 17$ of 1.

These operations can be executed within two different frameworks, each requiring a distinct implementation of the dictionary γ_A and register $\gamma_A[|\mathbf{a}\rangle]$:

1. Using an ordered register, each check, call, or insertion incurs a logarithmic time cost relative to the register size. In a worst-case scenario, there can be χ occurrences in the same register. Since it is reasonable for χ to scale at most polynomially in the number of qubits, checking the register $\gamma_A[|\mathbf{a}\rangle]$ has a cost scaling as $log(|\gamma_A[|\mathbf{a}\rangle]|) \sim log(N)$.

   The total time scaling would then be $O(\chi n_{\rho_A} log(N))$

2. Via Hash tables; each check, call, or insertion is $O(1)$ relative to the register size, although there is an additional cost for hashing the variable $\overline{\mathbf{a}}$. This scales linearly in the number of ones in the binary string $\overline{\mathbf{a}}$, which is at most N.

   The total time scaling would then at most $O(\chi n_{\rho_A} * N)$

In our implementation, we resorted to the second strategy, as by using the particle-hole formalism explained in Appendix 3.5, we only make use of binary strings with a fixed number of ones relative to the number of qubits, rendering the hashing cost a constant and bringing the total cost of the operation back to $O(\chi n_{\rho_A})$.

The memory scaling, making such heavy use of hash tables, depends greatly on different implementations, but it can be estimated roughly as a doubling of the weight of the starting wavefunction $|\psi\rangle$ due to the copying happening in the establishment of the hash table.

Mol.	Geometry (Å)	Basis-set	#Qubits	tracing method	χ
H_x chain	H —— H distance, 0.745	6-31G	$4 \cdot y$	Quadratic Direct	
H_2O	O 0.0 0.0 0.0 H 0.757 0.586 0.0 H -0.757 0.586 0.0	aug-ccpvtz [49, 79]	184	Direct	$1.2 \cdot 10^5$
C_6H_6	C —— H distance, 1.085 C —— C distance, 1.390	ccpvdz [49]	228	Direct	$2.5 \cdot 10^5$

Table 3.1: Summary of tested molecular systems together with the computational details. The parameter χ is the number of SD.

This method can be generalized to partial traces of mixed density matrices which cannot be written as (3.6.9), as for example a partially traced density matrix (not of a separable state), but the operation will have again a quadratic scaling in the dimension of the starting ρ.

3.7 Computational details

In the previous sections, we have introduced two main tools: encoding a sparse fermionic wavefunction into qubit space and using these wavefunctions to retrieve quantum information quantities, focusing on the partial trace of a density matrix.

We now want to present a practical analysis of our SparQ procedure to demonstrate the utility of the method. We use mutual information as the primary metric, defined by Eq. (1.10), as it has long been an essential tool for estimating correlations in quantum systems. Its estimation, however, requires the joint density matrix of the two subsystems whose correlation is under investigation.

The simulating framework is characterized by different packages and codes. The SparQ code implementing the trace (1.7) has been written in $C++$ to guarantee easy portability to parallelizing platforms such as OpenMP [42] while the sparse wavefunctions are all relative to CISD calculations found with the Pyscf [138] python library.
Coming to the actual simulated systems, we used multiple realizations of the hydrogen chain with a 6-31G basis or the analysis of the scaling of the tracing algorithm. To test the encoding in different mappings, we analyzed the water molecule (H_2O) with an aug-ccpvtz [49, 79] basis. At last, to test the full capabilities of the direct tracing algorithm, we studied how the entropy scales for different considered spaces for the benzene (C_6H_6) molecule expressed in the cc-pvdz [49] basis-set. All the data listed here is also summarized by Tab. 6.1.

3.8 Results

Scaling performance

To verify the theoretical scaling and to evaluate the actual computational cost, we use several hydrogen chains H_x of hydrogen atoms spaced 0.745Å apart to each other. Although a CISD description of such systems is not appropriate, we have chosen these simple systems only for scaling tests purposes.

The results are summarized in Fig. 3.1, showing the CPU time required to evaluate partial trace operation from N qubits to one qubit. Image a) shows the time requirements for the quadratic

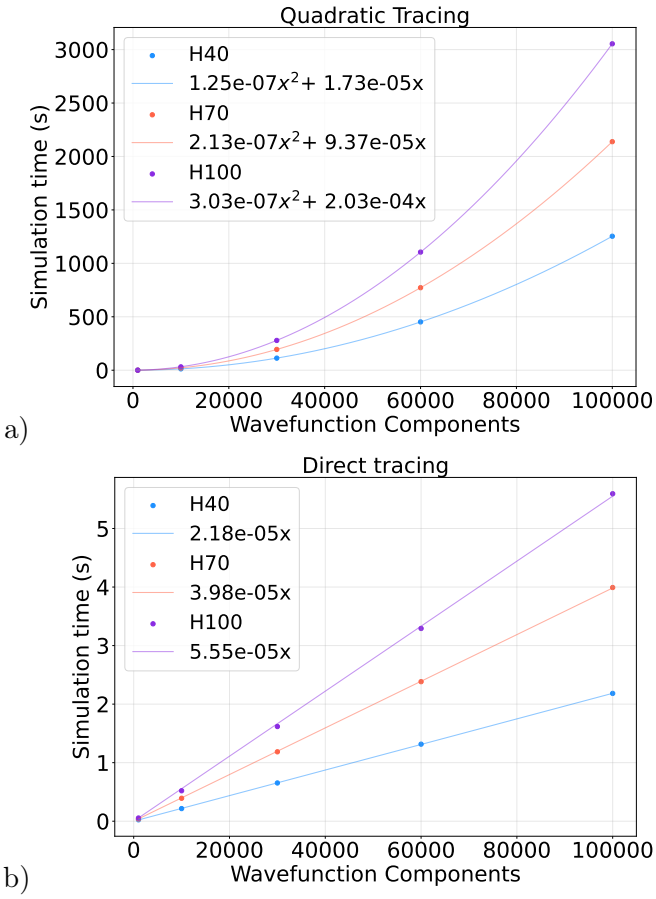
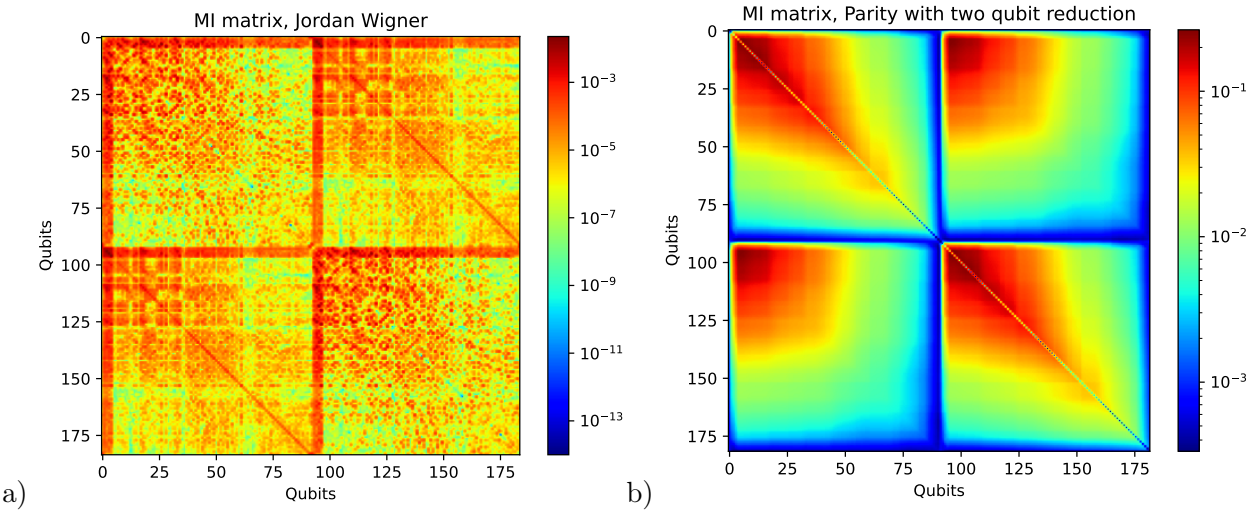


Figure 3.1: Time requirements for the partial trace from N qubits to one qubit over different values of number of components of the wavefunction a) Quadratic tracing method of Eq. (3.6.5), b) Tracing method of pseudocode 1 with linear scaling. As shown in 3.1 each chain H_x is expressed with $4y$ qubits.

scaling method described by Eq. (3.6.5), while image b) shows the scaling of the direct tracing method, as explained in section 3.6.4.

Looking at image 3.1-a), we can infer that the scaling of $O(\chi^2 N)$ for the (3.6.5) is achieved, as by fitting the parabolas not only we found a quadratic scaling for the χ parameter, but we also notice a linear scaling of the fitted coefficient of the quadratic term, coherently with the expected linear growth in N. This linear factor is also evident in 3.1-b), leading to the scaling of $O(\chi * N)$, as evaluated in section 26. This is because the shown simulation of 3.1 didn't make use of the particle-hole notation explained in appendix 3.5 specifically to show the behavior of the used Hash table.

This discussion leads us to believe that both Eq. (3.6.5) and algorithm 1 are valid methods to retrieve the partial trace and mutual information. However, while the quadratically scaling method of Eq. (3.6.5) can hardly handle up to approximately $\sim 10^5$ wavefunction components (3.1-a), the latter method, maintaining a linear computational cost as shown in 26, can likely be expanded to all the available components of a wavefunction. Its limitations are primarily due to the resources required for a binary representation of the wavefunction.



a) b)

Figure 3.2: Mutual information for the water molecule at CISD level. Jordan Wigner mapping is reported on the left panel, parity mapping with the two-qubit reduction on the right panel. The starting wavefunction for the mutual information was the simulation for the water molecule(H_2O) as detailed in Tab. 3.1.

Wavefunction in different mappings

The first proof of concept involves encoding the fermionic wavefunction using different fermion-to-qubit mappings, as illustrated in Fig. 3.2. We present the mutual information matrices $I(i,j) = I_{i,j}$ of the water molecule H_2O using two different encoding method for the wavefunction. To retrieve these figures in the parity mapping, we used the particle-hole transformation and CNOTs permutation between mapping, explained respectively in Appendix 3.5 and 3.4.

By using Fig. 3.2, we notice that the JW mapping (3.2-a) retains some chemical intuition, while the parity mapping (3.2-b) is of difficult interpretation since each qubit contains information of all the preceding ones.

Total entropy and active space

In the present section we would like to apply our wavefunction analysis to evaluate the total entropies of selected wavefunctions of the benzene molecule. In particular we are interested in comparing different selection methods of active spaces in CAS-SCF [90]. The choice of active space is not trivial, since some orbitals are more favorable than others in achieving lower energies during the orbital optimization. Several methods have been proposed to help with this task, of which we show the following:

- Fermi level selection. [110] The active space orbital selection occurs around the Fermi level of the HF occupation of the orbitals. Given its simplicity, it often misses the crucial orbitals if using the most restricted active space possible.

- Natural Orbitals Occupation Number (NOON) [81]. This method analyzes the occupation number of the Natural Orbitals (NO), which are the orbitals diagonalizing the one body reduced density matrix $\rho_{|\psi\rangle,i,j} = \langle\psi| a_i^\dagger a_j |\psi\rangle$ of a given fermionic wavefunction $|\psi\rangle$.

  The active space orbital selection is performed by including orbitals with occupation numbers that deviate the most from either empty occupation (0 electrons) or full occupation (2 electrons).

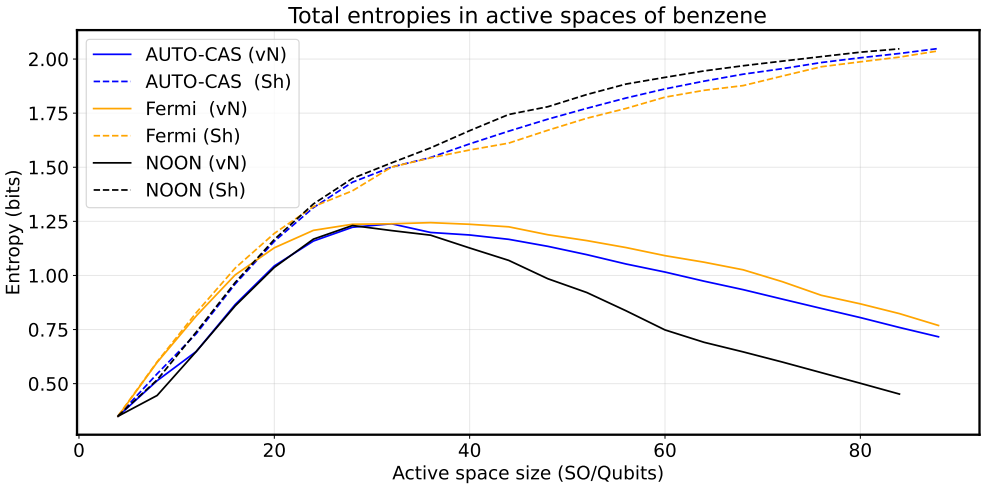


Figure 3.3: Total entropies as a function of the active space dimension and selection methods for a benzene molecule. The wavefunction analyzed is the CISD wavefunction of benzene with cc-pvdz basis. The Natural Orbitals for the natural orbitals occupation number (NOON) method were retrieved by an iterative procedure (Iterative Natural Orbitals) also based on the CISD wavefunction. Auto-CAS selects the orbitals maximizing single-orbital entropy, the Fermi method selects orbitals around the Fermi level of the HF orbitals. The continuous line is the von Neumann entropy of the whole space, while the dashed line is the Shannon entropy of the whole space of the state *measured* in the computational basis. Further details can be found in the main text.

This method has been proven to be quite effective, since the partial occupations are an indication of the presence of correlations in a CI expansion. However, since it relies on some chosen method of obtaining the wavefunction $|\psi\rangle$ and its related NO, it is not always applicable. Furthermore, one might also want to use different kinds of orbitals for the CAS-SCF, and not the NO.

- AUTO-CAS [136]. This method is based on an approximate DMRG calculation with low bond dimension. The selection of orbitals is then based on a maximization of the single orbital entropies (see Eq. (1.9)). This method plays a similar role than the NOON, since an orbital entropy greater than zero also means that the orbital is neither always occupied nor always empty. However, contrary to the NOON selection method, it can be applied to any orbital set.

Our goal is to analyze the behavior of the total entropy of active spaces of growing size, investigating differences between the above selecting methods.

Fig. 3.3 reports the total von Neumann and Shannon entropies as a function of number of orbitals/qubits included in the active spaces.
The system analyzed is the benzene molecule with a CISD wavefunction which is used in all steps of the calculation.

Looking firstly at the von Neumann entropy, we can see that even if the AUTO-CAS selects the most entropic orbitals, the resulting entropy of the total active space is always smaller than the one of the space selected around the Fermi level. While this might seem counter-intuitive, it can be explained by considering that the entropy of the total untraced density matrix $\rho = |\psi\rangle\langle\psi|$ must be zero for a pure state. Thus, the AUTO-CAS space is composed of the orbitals that allow to maintain, within the chosen space, the closest description to the original state $|\psi\rangle\langle\psi|$.

The same analysis has been performed using natural orbitals, to address the difference between the NOON selection method and AUTO-CAS in these orbitals. This comparison, however, reveals that there is a small difference, advocating for the similarity of the two methods. For this reason, in Fig. 3.3 we only show the NOON for clarity (This difference is shown in the supplementary materials). An interesting property of the simulation in natural orbitals is that they allow the smallest entropy at a fixed space dimension, evidence for greater effectiveness in describing the system with a restricted space with respect to the HF Canonical orbitals. This extends the already numerous studies regarding the convenient properties of the natural orbitals in quantum information [6, 96, 122] and quantum chemistry [23, 72].

One last interesting perspective is obtained by looking at the Shannon entropy of the space, which, contrary to the von Neumann one, is a monotone of the space dimension, as marginalization over a probabilistic variable can only decrease classical entropy.

3.9 Conclusions

In this work have introduced SparQ, a method to efficiently calculate quantum information theoretical quantities for sparse wavefunctions in quantum chemistry. The development of SparQ expands over the limitations of existing methods, mostly based on Density Matrix Renormalization Group (DMRG), which struggles with chemical systems exceeding a few hundred qubits due to the non-locality of the electronic Hamiltonian. This work therefore also addresses the limitations of studies based on transition operators, by generalizing the trace operator to a wide family of fermionic-to-qubit mappings and a great number of qubits.

By leveraging the sparsity inherent in many quantum chemistry wavefunctions, SparQ efficiently encodes fermionic wavefunctions into qubit space using fermion-to-qubit mappings like JW. This method enables the manipulation and analysis of large-scale quantum systems, significantly extending the reach of quantum information theoretical analysis. The practical utility of SparQ is shown through detailed procedures for the measurement of observables in the qubit space and partial trace of a variable number of qubits scaling linearly in the number of wavefunction components.

In practical applications, SparQ proves valuable in estimating correlations in quantum systems, as illustrated by the mutual information analysis of the water molecule and the entropy analysis of different possible active spaces of the benzene molecule. Overall, this demonstrated ability to handle large-scale wavefunctions opens new possibilities for understanding correlations and the behavior of complex molecular systems.

Future work will focus on refining fermion-to-qubit mappings, for example, by using the methods illustrated in the Appendices, to further reduce the computational overhead and expand the range of quantum information quantities that can be efficiently calculated using SparQ. The integration of various mapping techniques and the exploration of new application domains will continue to drive the evolution of quantum information analysis in Quantum Chemistry.

In summary, our work may contribute to the ongoing dialogue between quantum computing and quantum chemistry, paving the way for future innovations in hybrid quantum-classical computation and improving the understanding of correlations in quantum chemistry.

Chapter 4

Quantum Information reveals that orbital-wise correlation is essentially classical in Natural Orbitals

In this chapter, we explore the effect of orbital change through the lenses of Quantum Information Theory. Using Shannon and von Neumann mutual information, we find a stark difference between classical and quantum correlations with Hartree-Fock orbitals, while the Natural orbitals minimize this difference. We provide theoretical and computational justification and we discuss the potential impacts of this discovery. This suggests that, with the right orbital basis, wavefunction correlations are mainly classical. The methodology and results presented in this chapter are taken directly from the published (preprint) article in reference [97].

In recent years, there has been increasing interest in applying tools from quantum information theory to analyze chemical systems within the field of QCh [11, 12, 46, 47, 66]. This interest has led to a new understanding of known chemical properties [29, 48] and the development of efficient computational tools that focus computational efforts where they are most needed [21, 45, 136].

Part of this growing interest is likely caused by the recent encounter of QCh with quantum computing [157], a rapidly growing field that promises to bring great computational advantages to all computational sciences [17, 25, 31, 55, 71, 133] and in particular to QCh [10, 51, 64, 120]. Such an encounter, in addition to the direct application of quantum computing to QCh, promotes an exchange of ideas and technical tools in both directions, strengthening the synergy between quantum and classical computing systems.

This exchange has predominantly involved the transfer of concepts from QCh to quantum computing [19], such as, for instance, the Unitary Coupled Cluster (UCC) method, which adapts the concept of excitations from the Coupled Cluster method into a variational circuit suitable for quantum devices [14, 40, 53, 69, 84, 127, 151]. The analysis proposed here moves in the opposite direction, using typical quantum computing tools to better understand quantum chemical properties.

In the context of near-term quantum computing devices, the wavefunction ansätze inspired by quantum chemistry and used in the Variational Quantum Eigensolver require performances that are still far out of reach. For this reason, alternative methods based on heuristic variational circtuits are preferred because more effective, albeit the physical meaning of the wavefunction is often lost [76, 120, 123, 142]. Regardless of the chosen ansatz, different studies have shown that the

choice of the reference orbitals significantly impacts the computational efficiency [120, 135].

Within this context, Natural Orbitals (NO)4.1.1, have long played an important role related to the conjectured simplification in the expression of the wavefunction as hypothesized by Lödwin in 1956 [93]. Recently, Natural Orbitals have been also revealed to be a winning choice in empirical circuits for Quantum Computing, since they are the orbitals that optimize the energies of these ansatzes [96, 122]. Notwithstanding the amount of studies exploiting the NO, so far no direct numerical study has tackled the role played by this one-electron basis set on the quantum information properties.

With growing interest in quantum information theory, these analyses have been applied to understand entanglement and correlation in molecular systems[29, 47]. Interestingly, most of the existing works in the field are based on a Matrix Product State approach with analyses performed on Density Matrix Renormalization Group (DMRG)[129, 130, 136, 149, 150], calculations, whereas all Configuration Interaction (CI) methods are currently mostly ignored. In this chapter, making use of the SparQ [98] 3 method, we fill this gap for CI-based method for quantum information analysis. We employ an analysis with orbital correlation on configuration interaction wavefunction using the von Neumann entropy and Mutual Information, prominent tools in quantum information theory. Anticipating our result, we demonstrate, both mathematically and numerically, that classical and quantum correlations are almost identical only in the case we use Natural Orbitals as reference for the wavefunctions.

The chapter is divided as follows: referring to sections 1.1 and 4.1.1 for the introduction to quantum information theory and natural orbitals respectively, we delve into the core concept of this work in 4.1.2. Computational details are outlined in section 4.1.3, followed by results and discussion in section 4.2.

4.1 Methods and Theoretical Background

4.1.1 Natural Orbitals

While we gave an introduction of the NO in sec. 2.1.1, we deem necessary a further contextualization with regards to their role compared to other orbitals.

In the determinant-based approaches for the calculation of the electronic molecular wavefunction, the choice of the one-electron basis set is a key aspect. Indeed, a transformation of the MO basis leads to the modification of the SD ($|\mathbf{i}\rangle_{\mathcal{F}}$) coefficients in the expansion of the wavefunction Eq. (4.1.1), even considering methods that are invariant under a unitary transformation of the MO basis.

$$|\psi\rangle_{\mathcal{F}} = \sum_{\mathbf{i}=\{0,1\}^N} c_{\mathbf{i}} \, |\mathbf{i}\rangle_{\mathcal{F}} \tag{4.1.1}$$

Overall, the nature of the MO can have a marked effect on the numerical efficiency of the algorithm used to compute the wavefunction, such as the speed of convergence of the iterative Davidson procedure in CI methods [44].

For molecular systems qualitatively well described by a single SD, one can be tempted to base post-HF expansions on the Hartree-Fock Canonical Orbitals (HFCO) basis set, which are the eigenvectors of the Fock matrix. By contrast, the HFCO have shown to be a poor one-electron basis,

often exhibiting many non-negligible coefficients in (4.1.1) of similar magnitude in absolute value, indicating that a large number of SDs has to be considered to approach 100% of the wavefunction.

Since the beginning of quantum chemistry, an important effort has been devoted to defining a more effective orbital basis. Without climbing to be exhaustive, we cite here the Improved Virtual Orbital [141], the Modified Virtual Orbital [16], and the Internally Consistent SCF MO [44].

Within the quest of an optimal one-electron basis set, a peculiar role is played by the NO. As pointed out by Davidson [44] the virtual NO (those with an occupation number markedly lower than two) with the larger occupation numbers have the same extension as the occupied orbitals. It has been conjectured by Lödwin in 1956 [93] that the use of the NO minimizes the number of Slater determinants required to describe a large percentage of the wavefunction in the CI expansion. This compactness leads to important advantages. First of all, the length of the CI expansion can be reduced without a significant decrease of its quality. This can be done by removing SDs from the sum or by considering in the construction of the SDs only the NO with the largest occupation numbers. Moreover, even while keeping unchanged the length of the CI expansion, the calculation of the wavefunction (irrespective of its nature, CI, PT, or CC) is expected to be more effective. It is worth noticing here, that also the analysis is simplified when the wavefunction is more compact (for instance opening the way to the application of the effective or intermediate Hamiltonian formalism [26, 95]).

The construction of a CI wavefunction in its own natural orbital basis is not a trivial task, since the NO depend on the wavefunction which in turn is defined by the NO. In such a situation one can resort to an iterative procedure (Iterative Natural Orbitals (INO)) [73] which can be expected to converge to MO in which the wavefunction expresses a diagonal one-body reduced density matrix. Reaching self-consistency using this iterative method is rather expensive and in most case the procedure is stopped when a convergence criterion is reached.

On a side note, this iterative application is not always directly applicable, as in some cases the actual implementation of a method assumes that a specific MO basis is used, such as HFCO in MP2. For the sake of simplicity, for the following analysis, we only use CI methods, which in the most popular implementation allow the use of any one-electron basis as input. It is worth noticing, however, that some adaptations enable other methods to satisfy this requirement (e.g. MP2 [56]).

4.1.2 Two-qubit and one-qubit density matrices

It can be interesting to compare the approach here adopted with the formulation reported in the already mentioned works of Rissler *et al* [125] and Boguslawski [27], offering in this manner a direct comparison of mathematical relations with a purely computational approach following (1.7). More specifically, using these two past works we can establish some relations on the traced density matrices to one or two qubits. Let us first consider the two-qubit density matrices ρ_{ij} for which there can be two separate cases, i, j identify the occupation (through the JW mapping[75]) of either two spin orbitals with opposite spin, marked by $i_\uparrow, j_\downarrow$, or the two spin orbitals with the same spin $i_\uparrow, j_\uparrow$, where in this notation the arrow is only used to point out the spin of the linked spin orbital, while i, j can still vary from 0 to $N-1$. For the former case, we begin with Ref. [125], where it is shown that such a density matrix is bound to be diagonal, so that it is of the form:

Mol.	Geometry (Å)/source	Basis-set	#Qubits	*ab initio*-method	$\chi_{max,HFCO}$	$\chi_{max,NO}$
H_2O	O 0.0 0.0 0.0 H 0.757 0.586 0.0 H -0.757 0.586 0.0	aug-ccpvtz ccpvdz	184 48	CISD FCI	$1.2 \cdot 10^5$ $4.5 \cdot 10^8$	$1.6 \cdot 10^5$ 10^5
C_2	1.243 2.98 [37]	aug-ccpvtz	184	MRCI-CAS(8,8) CISD MRCI-CAS(8,8) CISD	$3.9 \cdot 10^5$ $1.7 \cdot 10^5$ $4.0 \cdot 10^5$ $1.6 \cdot 10^5$	$3.7 \cdot 10^5$ $1.4 \cdot 10^5$ $3.1 \cdot 10^5$ $1.6 \cdot 10^5$
Be_2	4.6 [80]	ccpv6z	560	MRCI-CAS(4,8) CISD	$2.0 \cdot 10^5$ $3.2 \cdot 10^5$	$7.7 \cdot 10^4$ $3.5 \cdot 10^5$
$(C_6H_6)_2$	[116] (s configuration)	ccpvdz	456	CISD	$8.3 \cdot 10^7$	$6.6 \cdot 10^7$

Table 4.1: Summary of tested molecular systems together with the computational details. The parameter χ_{max} is the number of SD in the variational space for the solving method, out of χ_{max} SD, we always took $min(\chi_{max}, 10^6)$ SD.

$$\rho_{i_\uparrow,j_\downarrow} = \begin{pmatrix} \rho_{|00\rangle\langle 00|} & 0 & 0 & 0 \\ 0 & \rho_{|01\rangle\langle 01|} & 0 & 0 \\ 0 & 0 & \rho_{|10\rangle\langle 10|} & 0 \\ 0 & 0 & 0 & \rho_{|11\rangle\langle 11|} \end{pmatrix} \tag{4.1.2}$$

This for us means exactly classical correlations since, as we already mentioned, eq (1.10) and eq (1.18) coincide for a diagonal ρ. This implies that the density matrix of whichever single qubit is also diagonal, as illustrated by (4.1.3).

$$\rho_i = tr(\rho_{i,j})_j = \begin{pmatrix} \rho_{|0\rangle\langle 0|} & 0 \\ 0 & \rho_{|1\rangle\langle 1|} \end{pmatrix} \tag{4.1.3}$$

It can then also be concluded that the entropy of a single qubit i necessarily obeys $S_{vN}(i) = S_{Sh}(i)$.
At this point, from the definition of mutual information between qubit i, j of Eq. (1.10), one has:

$$I_{vN}(i,j) = S_{Sh}(i) + S_{Sh}(j) - S_{vN}(i,j) \tag{4.1.4}$$

Furthermore, since $\rho_{i_\uparrow,j_\downarrow}$ is also diagonal as shown by Eq. (4.1.2), it also has purely classical correlations:

$$I_{vN}(i,j) = S_{Sh}(i) + S_{Sh}(j) - S_{Sh}(i,j) = I_{Sh}(i,j) \tag{4.1.5}$$

We have now remaining only the case of spin-orbitals with the same spin $i_\uparrow, j_\uparrow$, for which it is natural to ask if there are general diagonality conditions also for this last case. To this aim, starting from table 1 of ref [27, 28] we can write:

$$\rho_{i_\uparrow,j_\uparrow} = \begin{pmatrix} \rho_{|00\rangle\langle 00|} & 0 & 0 & 0 \\ 0 & \rho_{|01\rangle\langle 01|} & \rho_{|01\rangle\langle 10|} & 0 \\ 0 & \rho_{|10\rangle\langle 01|} & \rho_{|10\rangle\langle 10|} & 0 \\ 0 & 0 & 0 & \rho_{|11\rangle\langle 11|} \end{pmatrix} \tag{4.1.6}$$

In this expression, components are defined in terms of multi-body density matrices. In particular, the *off-diagonal* term $\rho_{|01\rangle\langle 10|} \equiv \rho_{|10\rangle\langle 01|}$ for $\rho_{i_\uparrow, j_\uparrow}$ is computed as:

$$\rho_{|01\rangle\langle 10|} = {}_{\mathcal{F}}\langle\psi|(-a_j^\dagger a_i + a_j^\dagger a_{\bar{i}}^\dagger a_{\bar{i}} a_i + a_j^\dagger a_{\bar{j}}^\dagger a_{\bar{j}} a_i - a_j^\dagger a_{\bar{j}}^\dagger a_{\bar{i}}^\dagger a_{\bar{i}} a_{\bar{j}} a_i)\,|\psi\rangle_{\mathcal{F}} \tag{4.1.7}$$

where the barred index $\bar{j}$ is used to indicate spin-down orbitals indices and the expectation value takes place in the Fock space as in (2.30)[1].
In this expression the largest contribution is generally expected to come from the first term, $a_j^\dagger a_i$, representing a looser condition than the second term for nullity of the expectation value between SDs. For this reason, it is expected that decreasing (setting to zero) its value will likely bring to a more classical correlation pattern by bringing $\rho_{i,j}$ in a diagonal shape, and this is precisely the role of the NO.

4.1.3 Computational Details

The molecules considered in this work were chosen to encompass systems with different kinds of chemical interactions and electronic correlation, such as vdW interactions, single and multiple bonds as reported in table 4.1. We used different quantum chemistry methods, depending on which is best suited to tackle each molecule.

All systems were first treated using a specific post HF method in the restricted-HF regime. In the MRCISD method the wavefunction is obtained diagonalizing the Hamiltonian on the space of all possible singles and doubles excitations out of the most important SD of the CAS(e,N) solution. [2] As for the CASSCF part of this algorithm, it was only fully implemented for the simulation aiming at the natural orbitals, leaving aside orbital optimization (CASCI) for the simulation labeled by the HFCO

We use correlation consistent basis sets [49], using the aug-ccpvtz [79] for H_2O and C_2, the ccpv6z basis for Be_2 [80] [114, 118, 153] and ccpvdz [116] basis set for $(C_6H_6)_2$. All basis set descriptions were taken from `basissetexchange.org` [119].

We used different packages and codes for the simulations and data analysis, including the Pyscf [138] python library, the Orca Package [107] for MRCI calculations and the SparQ [98] code mainly for the implementation of the trace (1.7) which has been implemented in-house in C++ to guarantee easy portability to parallelizing platforms such as OpenMP [42].

For all the systems we report the quantum and classical entropy and pair-wise mutual information for the spin-orbitals. We undertake this analysis for two different sets of orbitals for each system, the HFCO and the NO (computed in the iterative scheme). The analysis was carried out using the first χ SD with the largest coefficient (in absolute value) in the expansion (4.1.1). The chosen value of χ was however large (10^6), and in most cases includes all the SD of the configurational space of the wavefunction. In Table 4.1 it is also reported the χ_{max} value, representing all possible SD of non-zero coefficient.

[1] Technically, both the ladder operators and $|\psi\rangle$ are assumed to lye in the separable qubit space by [125], however this expectation is invariant under the fermionic-to-qubit mapping isometry.

[2] Actually, the FCI in CAS is made at configuration state function (CSF) level with a selection threshold for inclusion in MRCI of $\tau = 10^{-5}$.

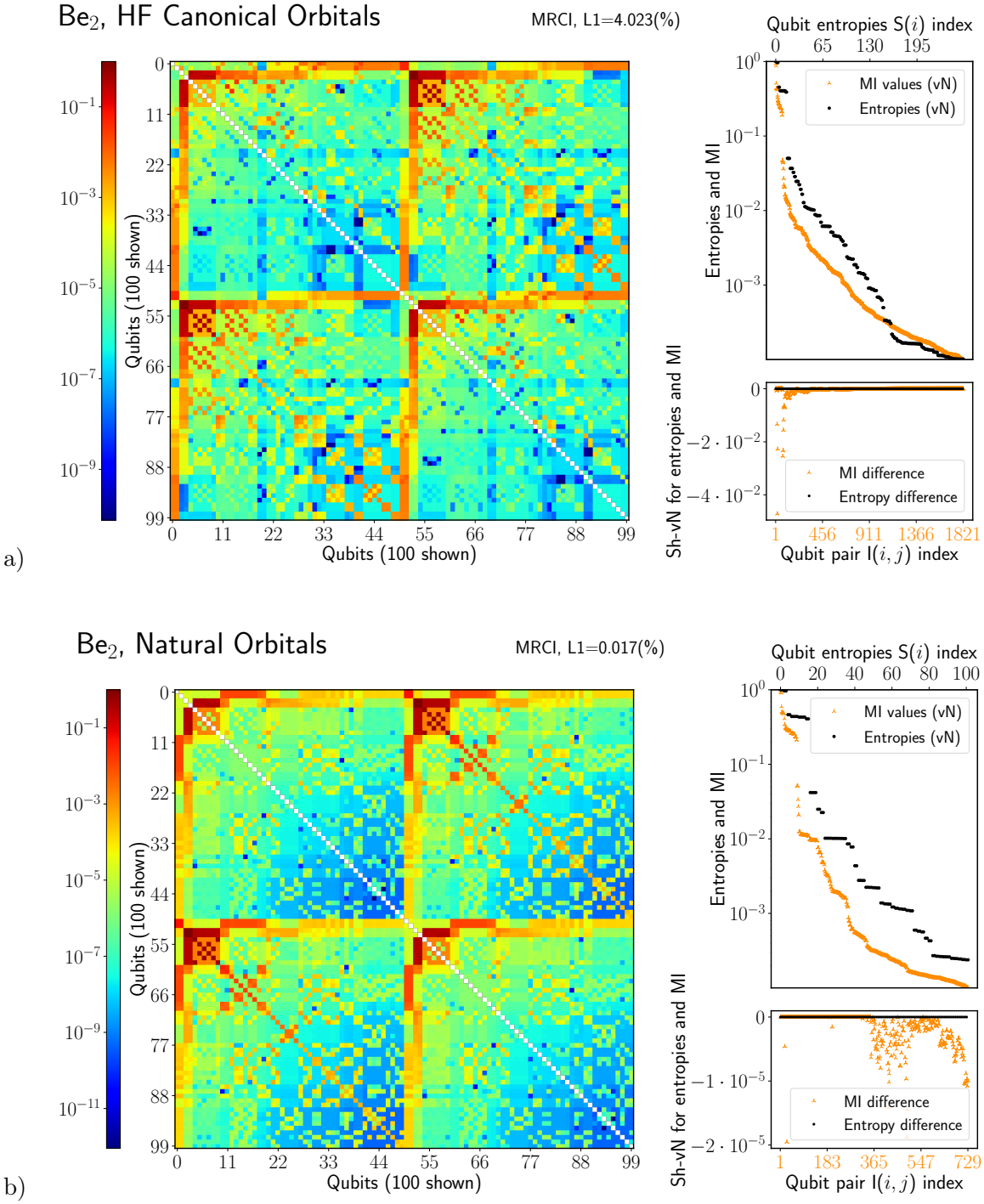
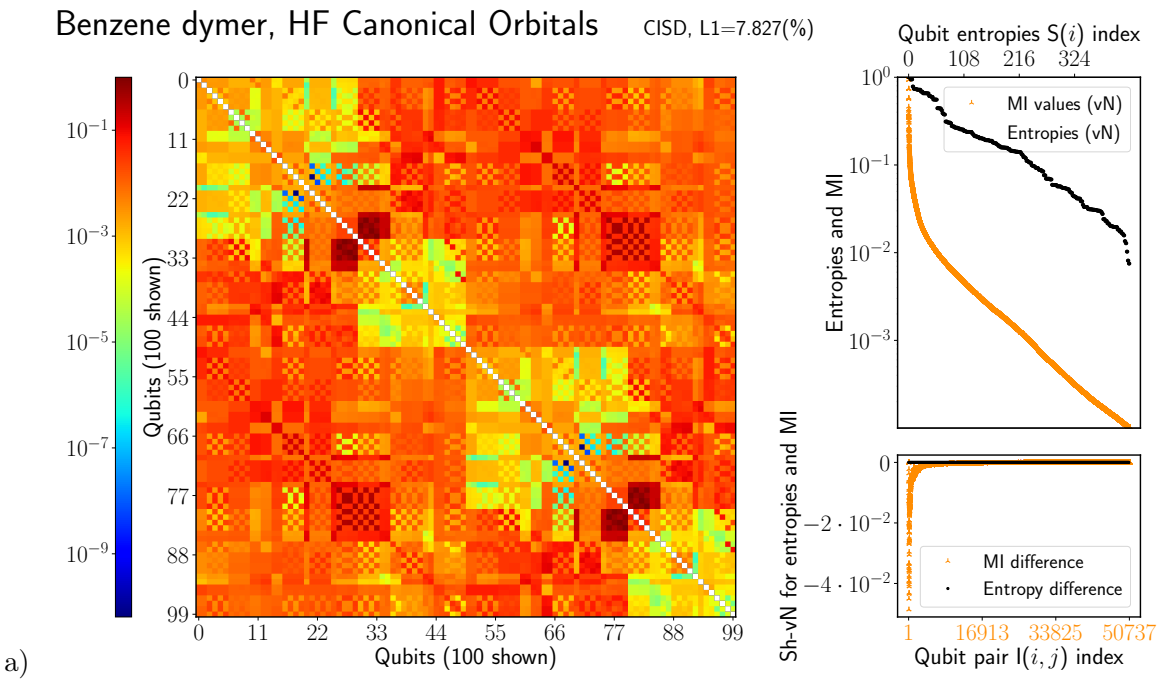
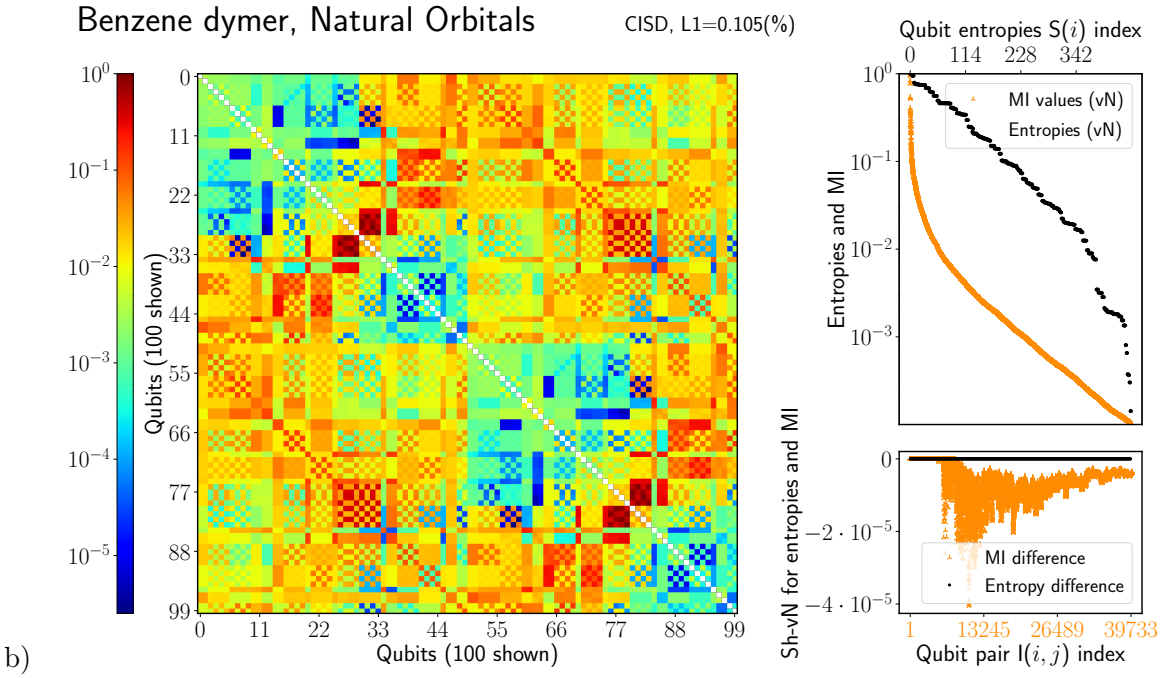


Figure 4.1: Classical and quantum orbital correlation for the MRCI wavefunction of Be_2 molecule expressed in Hartree-Fock Canonical Orbitals (top panel) and Natura Orbitals (bottom panel). *Left*: Heatmap of the von Neumann mutual information $I_{vN}(i,j)$ values in the upper triangle of the matrix and of $I_{Sh}(i,j)$ in the lower triangle. To improve visualization we only displayed the 100 qubits/spin-orbitals with the largest entropy. *Right top*: Plot of sorted elements $sort(I_{vN})$ of the $I_{vN}(i,j)$ matrix (red empty circles) and sorted single qubit entropies $sort(S(i,j))$(black filled circles); only values above a certain treshold are diplayed. *Right bottom*: Difference of $sort(I_{Sh}) - sort(I)_{vN}$(orange) and corresponding differences between sorted single qubit entropies.



a)



b)

Figure 4.2: The two images report the result with the HFCO and the NO one-electron basis sets for the $(C_6H_6)_2$ system. *Left*: Representation of the von Neumann mutual information $I_{vN}(i,j)$ values in the upper triangle of the matrix and of $I_{Sh}(i,j)$ in the lower triangle. i,j are limited to the 100 qubits with the largest entropy. *Right top*: Descending trend of $I_{vN}(i,j)$ matrix $\forall i,j$ (dark red) and same trend for the single qubit entropy $S_{vN}(i)$(black), both quantities are shown up to a 10^{-4} threshold. *Right bottom*: Difference of $sort(I_{Sh}) - sort(I)$(orange) $\forall i,j$ and same comparison for all the single qubit entropies.

4.2 Results and Discussion

For our analysis of the quantum and classical information properties of the quantum chemistry's wavefunctions we have chosen molecular systems with different kinds of bonds and interactions, to investigate whether our findings might have a general relevance.

We carried out a separate analysis for each system, delving into a comparison of the relevant quantities for our analysis. System description and computational details are summarized in table 4.1. We begin to illustrate our results by focusing on figures 4.1 and 4.2 that contain two panels, the upper one referred to MRCI calculations on Be_2 at equilibrium distance using the Hartree-Fock Canonical Orbitals, the bottom one referred to the same calculation using Natural Orbitals. Within each panel, on the left is represented a heatmap plot of the quantum and classical mutual information matrix, showing for each qubit/spin-orbital index i, j the values of the matrices, with the indices running over the 100 qubits with the largest entropy, of which the first 50 are linked to (via Jordan Wigner mapping) spin-up orbitals and the second half to spin-down orbitals. Exploiting the symmetry of the Mutual Information matrix we represented on the same heatmap both classical and quantum matrices. On the upper (top-right) triangular matrix, we have the von Neumann mutual information $I_{vN}(i,j)$ defined in (1.10), whereas the lower (bottom-left) triangular matrix reports the Shannon Mutual Information $I_{Sh}(i,j)$ defined in (1.18). This representation allows one to visually compare the differences between the two classical and quantum mutual information. A more detailed analysis is shown in the plots on the right side of the panels. Here, the upper part of the figure shows the diminishing sorted array $sort(I_{vN,\rho})(i,j)$ and single qubit entropies $sort(S_{vN,\rho})(i)$ for each (i,j) pair and i, showing only values up to 10^{-4}. Finally, in the lower plot of the panel, we represent the quantity $sort(I_{Sh,\rho}) - sort(I_{vN,\rho})$ with the sorting running over indices i, j, thus the subtraction between the sorted array of $I_{Sh,\rho}$ and the curve on the top plot. A similar analysis is also reported for entropy, thus showing the difference of $sort(S_{Sh,\rho}) - sort(S_{vN,\rho})$ with the sorting running over the qubit index i.

Figure 2 shows the same data for the benzene dimer example. Both series of data in Figure 4.1 and 4.2 showdifferences of the structure of the von Neumann mutual information between the HFCO and the NO, namely the differences between the upper-right triangular part of the heatmap and the lower-left triangular part. It is already clear visually that the differences are significantly attenuated in the lower panels, where the wavefunctions are expressed in terms of Natural Orbitals. This is in line to what we already discussed in section 4.1.1, since the NO are supposed to provide, in general, a basis set in which the wavefunction is represente in a more compact structure.

The difference between the von-Neumann and Shannon quantities is larger in the HFCO basis than in their NO counterparts, as visually perceptible in Figure 4.1(a) and 4.2(a). Numerically this effect can be estimated numerically using the following expression:

$$L1_\rho = \frac{\sum_{i,j>i}^{N} I_{vN}(i,j) - I_{Sh}(i,j)}{\sum_{i,j>i}^{N} I_{vN}(i,j)} * 100 \tag{4.2.1}$$

The values of the $L1_\rho$ parameter for the different molecules using both the HFCO and the NO are reported in Tab. 4.2.

The $L1_\rho$ difference (in percentage) is a quantity always greater or equal to zero because of the inequality (1.20). We reported $L1_\rho$ both in the caption of Figures 4.1 and 4.2 in table 4.1 as

Mol.	Div. summary	$E_{HFCO}(E_h)$	$E_{NO}(E_h)$	$L1_\rho$ (%) HFCO	$L1_\rho$ (%) NO	γ_{HF}	γ_{NO}
H_2O	aug-ccpvtz	-76.337	-76.336	20.942	0.058	0.132	$\sim 10^{-8}$
	ccpvdz (FCI)	-76.244	-76.244	5.0508	0.060	0.260	$\sim 10^{-9}$
C_2	1.243Å MRCI	-75.764	-75.780	7.735	0.186	//	//
	1.243Å CISD	-75.737	-75.733	16.753	0.011	0.150	$\sim 10^{-6}$
	2.98Å MRCI	-75.548	-75.561	3.002	0.112	//	//
	2.98Å CISD	-75.433	-75.428	10.857	0.006	0.147	$\sim 10^{-5}$
Be_2	MRCI	-29.307	-29.306	4.023	0.017	//	//
	CISD	-29.293	-29.293	18.587	0.001	0.127	$\sim 10^{-5}$
$(C_6H_6)_2$		-462.627	-462.627	7.827	0.105	0.023	$\sim 10^{-4}$

Table 4.2: Summary of results for tested systems, the second column reports the division as shown in tab. 4.1. $L1_\rho$ in eq (4.2.1) has been introduced as percentage difference in correlations between (1.10) and (1.18) (see Eq. (4.2.1)).The γ parameter is introduced in Eq. (4.2.2).

percentage. The reported computed data clearly shows that Natural Orbitals one-particle basis set gives a description of the correlation which is essentially classical, drastically reducing the $L1_\rho$ difference between the von Neumann and the Shannon mutual information.

It is worth noticing that all the above analyses are based on approximate wavefunctions, and, moreover, that the Natural Orbitals are also approximate because of two different reasons. Natural orbitals are indeed obtained by an iterative procedure, namely, they are Iterative Natural Orbitals. The approximations comes therefore from two distinct sources: the first is the order of the iteration, the second is due to the fact that the ground state method is also approximated by the post Hartree-Fock method used in the specific case. The convergence as a function of the number of iterations can be followed by monitoring a parameter that quantifies the diagonal shape of the density matrix at each iteration :

$$\gamma = \frac{1}{N_e} \sum_{i,j,i\neq j}^{N} |\rho_{i,j}| \tag{4.2.2}$$

with N_e being the number of electrons (and therefore also the trace of the one-body reduced density matrix).

We report in Figure 4.3 how this quantity changes along the iterative procedure. It is clear that the NO procedure properly converges. It is worth noticing that during the process the energy can also increase since the procedure is not variational, an example is shown in tab 4.2 by the molecules H_2O, C_2 and Be_2, however, moving from the HFCO to the INO basis the energy increase is limited to a few hundredths of a Hartree, at most. On the other hand, we noted that stopping the iterations at the first step rarely increases the energy, but γ can still display quite a large value, as clearly shown in Figure 4.3.

Since we are interested in reducing the off-diagonal terms of the density matrix, which are sensitive to the accuracy of the NO, to guarantee to use properly converged natural orbitals, we carried on the iterative procedure until either γ was lower than $\sim 10^{-8}$, or a maximum of about ~ 10 iterations were reached. The value of γ at the last iteration is also reported in table 4.2, from which one can highlight the marked difference of the γ value of the HFCO and INO.

The convergence of γ and the corresponding convergence of the $L1_\rho$ values are reported in

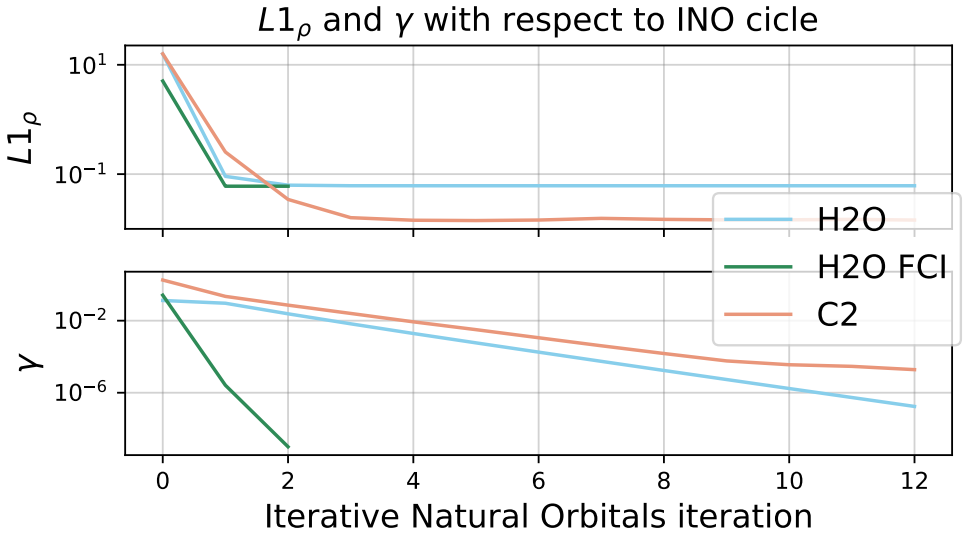


Figure 4.3: Improvement of iterative natural orbital through different iterations via γ and the effect on the resulting L1. H2O and H2O FCI refer to the first and second line respectively of the H_2O molecule in Tab. 4.1, the same is meant for C2 but without the ambiguity of different basis set and P-HF method of simulation.

figure 4.3.

In the previous section, we have reported a solid evidence for the classical behavior of the orbital-wise correlation when the wavefunction is expressed in the basis of natural orbitals. For example, the heatmap plots of figures 4.1 and 4.2 are a clear visual representation of the effect of the one-electron basis set on the von Neumann mutual information and its relation with the correspondin classical mutual information based on the Shannon entropy. The $L1_\rho$ value, which is providing an estimate of the differences between classical and quantum correlation, is significantly smaller in Natural Orbitals with respect to its value in the Hartee-Fock basis (up to 100 times in some cases). When sorted, the values of mutual information (red spots in the right panels in the figure) decrese much faster using NO than with HFCO. This result has been already pointed out by us recently [122] but here is tested on a different range of systems and larger molecular sizes.

From the computational point of view, the classical behavior of correlations in the NO in principles basis paves the way for a simplification of the computational resources required to attain a given precision in the simulation. Two important implications arise. Firstly, these findings show, once more, the importance of the choice of orbital basis for CI-based method. In this regard, the use of NO in classically computed quantum chemistry is strongly suggested whenever possible since their key role in CI calculation has been longly recognized. However, we believe that their potential should not be limited to such methods. In Quantum Computing, we have shown how the use of NO can be beneficial for the efficiency of variational algorithms [96, 122], and the analysis reported in this work confirms these considerations. Because these properties, it is clear that Natural Orbitals should always be preferred to properly analyze the wavefunction in terms of orbital-correlation. Indeed, other orbital choices would likely overestimate the quantum contribution to the mutual information.

The second implication delves into the foundations of quantum computing. If correlations in quantum chemistry are predominantly classical in nature, additional attention is warranted to the

theoretical need for a quantum computer to solve the electronic structure problem. In line with these concerns, other authors have expressed skepticism and reservations about the actual potential of quantum computing in quantum chemistry. [88].

Concerning the classical structure of the von Neumann mutual information when the wave-function is built on NO, the results here reported are consistent with the discussion developed in section 4.1.2 on the relationships described in (4.1.7) and based on chemist's multi-bodies reduced density matrices to determine the traced density matrices required to calculate the von Neumann mutual information. Not only we confirm what is described by Eq. (4.1.7), but also validate all the contributing factors that led to this conclusion. Specifically, the entropies of single qubits are characterized entirely by classical entropies ($S_{vN} = S_{Sh}$), as shown by top right of figures 4.1 and 4.2. Additionally, the mutual information between two qubits, when mapped to orbitals with opposite spins, also exhibits a purely classical characterization, independently on the spin orbital basis. For the remaining combination of same-spin qubits, some residual quantum contribution to the mutual information is remaining, even if this is significantly reduced (up to 100 times) in the fully converged INO basis. This residual, which is still present with fully converged INOs, Considering one body reduced density matrix diagonal, the orginin of this residual L_1 difference might be due to the higher order multi-body reduced density matrices, as pointed out by quation (4.1.7).

4.2.1 Conclusions

In summary, this chapter investigated whether the orbital-wise correlations in Quantum Chemistry are quantum or classical in nature, employing techniques from Quantum Information Theory. We conducted numerical experiments on moderate-sized systems, which vary in chemical nature and types of bonding. These systems were treated using high-level post-Hartree-Fock methods and involved up to approximately 550 spin-orbitals. Our numerical tests, supported by mathematical reasoning, illustrate how Natural Orbitals significantly reduce the purely quantum von Neumann-only qubit-qubit (spin orbital-spin orbital) correlations, predominantly leaving only the classical contributions. The computational and theoretical evidence obtained in this study lays the foundation for computational simplification in the processing of quantum information quantities in quantum chemistry. This provides additional mathematical justification for the more widespread use of Natural Orbitals and raises further questions about the actual computational complexity of the multi-body problem in quantum chemistry.

Chapter 5

QIDA, Quantum Information Driven Ansatz

This chapter proposes a new hardware-efficient empirical Ansatz for the Variational Quantum Eigensolver (VQE) in Quantum Chemistry, designed using quantum mutual information derived from classical Quantum Chemistry methods. The idea is to initially employ Møller-Plesset (MP2) 2.3.2 perturbation theory to generate approximate Natural Orbitals (NO), which are ideal for developing compact wavefunctions. We then evaluate the quantum mutual information matrix to design entangling blocks for the quantum circuit, resulting in short-depth efficient ansätze. This chapter is based on the published work of Ref.[96].

As we introduced in Sec. 1.2.3, the widely known VQE algorithm is based on different approaches for building the variational quantum circuit. Among all of them, two major classes can be found. The first consists in translating classical Quantum Chemistry approaches [1] into the language of quantum computation. One of the most representative examples is the Unitary Coupled Cluster (UCC) method[14, 40, 53, 68, 84, 126, 151] which is a unitary implementation of the widely used Coupled Cluster classical method. The second approach, on the contrary, starts from wavefunctions that are directly built by exploiting the characteristics of the quantum hardware (for instance, by reducing as much as possible the depth of the quantum circuit), regardless of connections with the Chemistry of the problem. This empirical approach defines the so-called Heuristic Ansatz[76, 120, 123, 142], which is composed of repetitions of blocks of rotations and entanglements, set up without any information of the chemical system. Compared to the former approach, this one better exploits the capabilities of the quantum hardware, at the price of losing the chemical meaning of the variational ansatz.

This chapter aims to join these two approaches by building, within the realm of the latter method, short-depth circuits integrating knowledge derived by Classical albeit affordable Quantum Chemistry calculations such as second order Møller-Plesset perturbation theory (MP2). [106] in doing so, we provide an effective starting point for latter VQE runs. The key physical quantity for this bridge is the Quantum Mutual Information (QMI). In classically computed Quantum Chemistry, the Quantum Mutual Information has already appeared under different aspects, its main objective however has always been to give a measure of correlation between different orbitals. [47] During the years, this tool has been used to help converge DMRG calculations [89, 125] or to establish

[1] in the sense of not arising from quantum computing

a black-box method of selecting active orbitals within the CASSCF framework. [136] In quantum computing, QMI has already been used for suggesting the best logical-to-physical assignment for the qubits [145] or improving existing algorithms. [159]

In our approach, the Quantum Information Driven Ansatz (QIDA) is built with a circuit topology that reflects the QMI derived by preliminary quantum chemistry calculations as will be detailed in the next sections. This ansatz represents a compact and affordable starting point on which to construct more complex wavefunctions. Together with such analysis, a further enhancement of the results offered by our scheme will be also provided by the use of Natural Orbitals (NO) as the preferred basis where developing correlated compact wavefunctions, as recently shown by us in the context of WAHTOR algorithm. [120, 121] The background of quantum mutual information-aided VQE makes this work complementary to the approach reported in ref. [159], which treats a similar aim in the context of a UCC ansatz within an adaptive framework.[64]

The chapter is organized as follows: in section 5.1, we introduce our scheme for obtaining ansätze. Once the method is clear, we present in section 5.3 our results and we conclude the work in section 5.4.

5.1 Method

5.1.1 The scheme

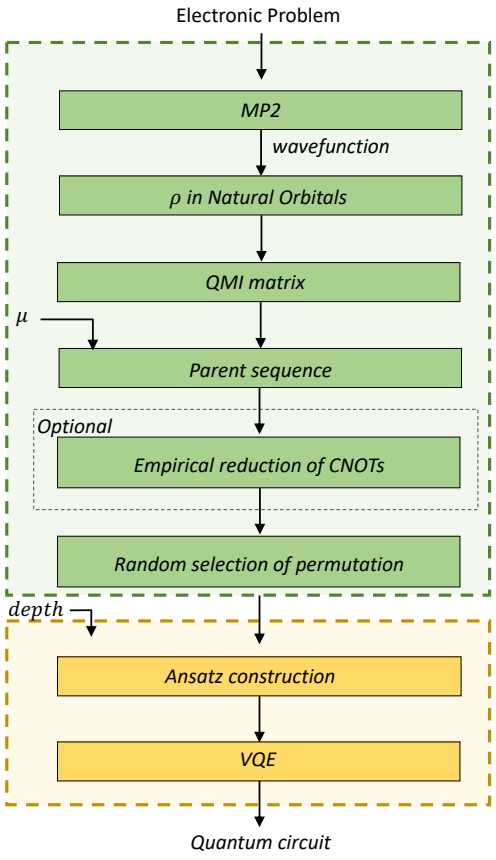


Figure 5.2: Description of the proposed QIDA scheme to obtain an ansatz from information inferred by a post-Hartree-Fock method, MP2 in this work.

Resource-efficient classical quantum chemistry methods provide valuable information for quantum computing algorithms addressing electronic problems. This is not a new concept as indeed the starting point for VQE calculations is often the HFCO basis, even if starting from some other basis is possible. In this work, we have developed a scheme that uses quantum mutual information derived from classical calculations to design entangling blocks for hardware-efficient ansätze.

As a first step, we perform a Post-HF (PHF) calculation to obtain an approximation of the GS $|\psi_{PHF,GS}\rangle$. The choice of the method is usually dictated by the size of the problem. We perform a second-order Møller-Plesset (MP2)[106] calculation which scales with $O(n^5)$[11] in the number of orbitals. We stress an important point about these wavefunctions, if one uses, as done here, MP2 (or CCSD) wavefunctions, the so-retrieved state will not have the energy given by the MP2 calculation, as perturbation theory is intrinsically not variational. For this reason, a comparison between the MP2 state and its classically retrieved energy is of limited use. Nevertheless, right from our preliminary tests the found MP2 wavefunction proved to be good enough for our purposes. As

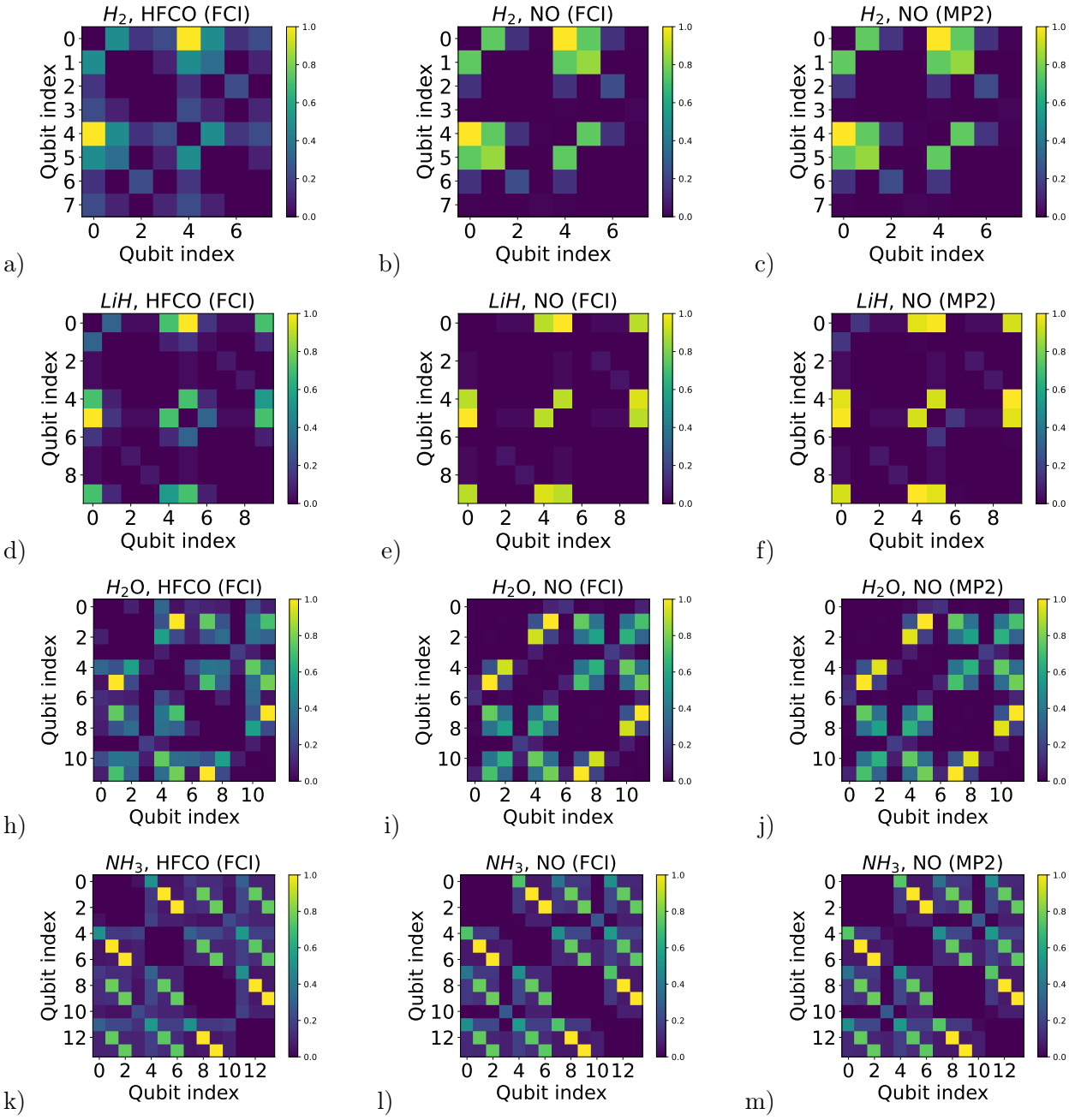


Figure 5.1: Representation of the quantum mutual information matrices $I(i,j)$ for qubits mapping the spin-orbitals of the molecular systems considered. The rows span over the molecules, while the columns span over the orbitals and the states: starting from the left we meet FCI state in HFCO for the first column, and then FCI and MP2 states each in its respectiveNOfor the second and third column.

a following step, we map [34, 35, 75] the state $|\psi_{PHF,GS}\rangle$ on a qubit register to obtain $|\psi_{q,PHF,GS}\rangle$, by encoding each excitation that contributes to each Slater determinant of $|\psi_{PHF,GS}\rangle$.

At this point, we find the corresponding I matrix by using equation 1.10. Some explicit examples of these matrices are given in fig 5.1, with elements ranging from 0 to 1 because of enforced normalization.
Overall, we can consider these values of I as a measure of correlation [158], without discerning between the classical correlation and the quantum correlation.
The aim of this work is to explore the general concept of constructing variational ansätze by exploiting the correlation pointed out by the mutual information matrix I. We are confident that we will not be able to capture all electron correlation but to catch the major contribution in a maximally compact circuit ansatz. Within this framework, we can therefore identify which pairs have a larger contribution to this correlation by fixing a threshold $\mu \in [0, 1)$, dividing our matrix into two separate sets of the pairs falling either above or below μ. In a more general case, different thresholds might be chosen to separate this interval in more sections or additional circuits may be added to this starting ansatz to reach the desired level of description. However, in the present work this single selection rule is used to design our empirical ansatz, by considering the entangling block of the VQE to be composed of CNOTs only between qubits i, j for which $I(i, j) > \mu$. In this framework, coherently to usual heuristic approaches [120], the CNOT is chosen as an entangler (i.e. correlator) between qubits. Of course, other two-qubit quantum gates could also fulfill the same role.

For our analysis based on a single threshold μ, the general idea is that a large value selects the most relevant correlations between qubits. By decreasing the μ parameter, one can now choose in a consistent way how many CNOTs to include, to build more and more correlated ansätze, at the price of a longer circuit depth. Clearly, a very small value of μ could include all qubit couples, coming back to a complete ladder-like ansatz, but this is not the aim of the present work, which is to build a first short ansatz including only a few qubits and CNOTs, recovering partially the variational energy. To further reduce the number of CNOTs of the ansatz we also introduce a protocol that will be described in the next subsection. Eventually, after having chosen a μ and having applied a CNOTs reduction method, one will have selected a list of CNOTs to be applied as an ansatz, named the *parent* sequence.

Since our algorithm doesn't specify the order of the involved CNOTs, and considering that one has to entangle at least $\mathrm{O}(n)$ qubits in a n qubit register to achieve meaningful results, the number of possible orderings will grow as $n!$. This fact, which is an evident computational bottleneck, does not in practice preclude the method from correct functioning, since n can be kept small by choosing an appropriate value of μ. Anticipating our results, we will also show that the choice random permutation between the $n!$, will not be a relevant factor with multiple ansatz repetitions and will give anyway better statistical results than the common ladder-entangler block with the same numbers of CNOTs. In this context, hereon the role of the sequence given as output, defined *parent* sequence, will be to represent the set of all possible permutations generated from itself.

Empirical reduction

In order to retain the scaling of the two-qubit gate count of the ladder ansatz, bilinear in the number of qubits and depth, we found that it is convenient to remove some of the CNOTs belongign to groups which are cross-entangled, i.e. forming closed-loops of connections between different qubits. Figure 5.3 can help to clarify the concept of cross-entanglement. In this specific case all the pairs obtained from qubits 1,5,7 have an I between each other greater than a certain μ threshold. As described

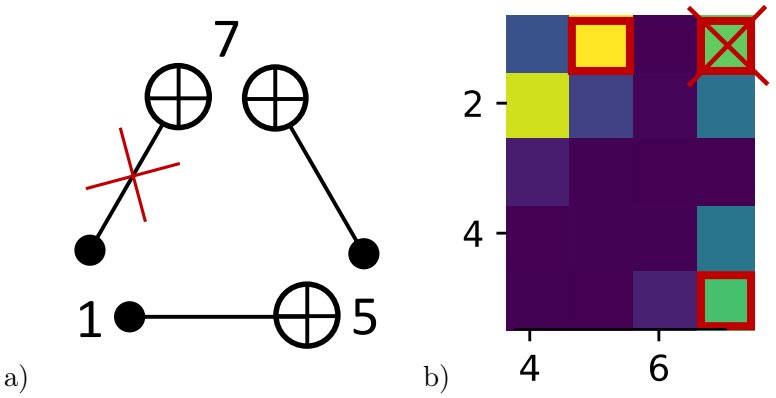


Figure 5.3: Cross-entanglement reduction. All numbers refer to the qubit indices. a) We have three CNOTs but we do not consider the one connecting the qubits 1 and 7. b) Representation on I map of the reduction. The red frame identifies the pairs that we connect with a CNOT on the circuit whereas the crossed one is left out.

in the previous section, our procedure suggests to insert three CNOTs corresponding to all three pairs. Nevertheless, in a variational algorithm having multiple depths, it should be sufficient to connect these qubits only once. This means that it is possible to remove one of the three CNOTs, as represented in figure 5.3[a]. We represent graphically this operation by putting a red frame on the considered spots and erasing the one that will not appear on the circuit, as done in figure 5.3[b]. Starting from this consideration, we have implemented an empirical strategy in which only the first spot from each line of the upper triangular part of the I matrix such that $I(i,j) > \mu$ is selected. As an example, figure 5.4 shows the pairs taken into account for the H_2O and NH_3 molecules.

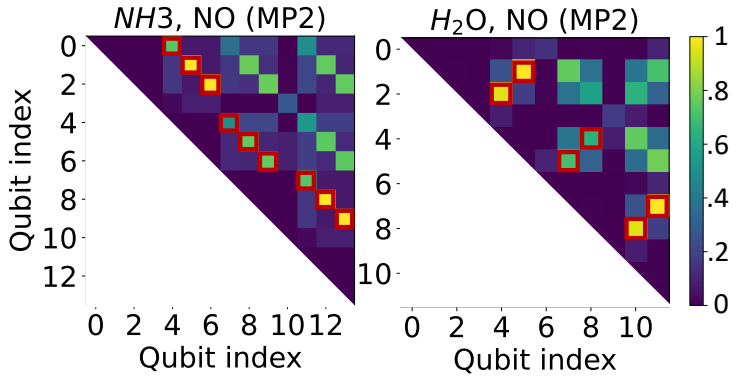


Figure 5.4: Empirical reduction criterion applied to H_2O and NH_3 molecules. As described in the main text, spots with red frames identify the qubits pairs that will be entangled with the two-qubit gate (CNOT in our case).

Mol.	Coordinates (Å)	Qubits	$E_{ex,c(\frac{Kcal}{mol})}$	μ	# elements in parent sequence
H_2	H 0.0 0.0 0.0 H 0.0 0.0 0.74	8	15.636	0.5	6
LiH	Li 0.0 0.0 0.0 H 0.0 0.0 1.595	10	12.6446	0.5	6
H_2O	O 0.0 0.0 0.0 H 0.757 0.586 0.0 H -0.757 0.586 0.0	12	31.006	0.5 0.7 0.5 *red.*	12 9 6
NH_3	N 0.0 0.0 0.1211 H 0.0 0.9306 -0.2826 H 0.8059 -0.4653 -0.2826 H -0.8059 -0.4653 -0.2826	14	41.287	0.5 *red.*	9

Table 5.1: Summary of data of the systems under analysis and computational details. H_2 has been analyzed with 631-G basis while every other molecule used STO-3G basis set and frozen core approximation at the HF level

5.2 Computational details

5.2.1 Tested Systems

The part of our algorithm that uses classical Quantum Chemistry approaches, such as Hartree-Fock (HF) and electronic integral calculations, is performed using the PySCF[138] Python package, whereas the MP2 wavefunctions are obtained through an adaptation of the MolCAs package.[9]

The fraction of the correlation energy ϵ is defined by

$$\epsilon = \frac{E_X - E_{HF}}{E_{ex,c}}, \quad E_{ex,c} = E_{FCI} - E_{HF} \tag{5.2.1}$$

where E_X is the energy obtained from the variational calculation and E_{HF} and E_{FCI} are the Hartree-Fock and full configuration interaction energies.

The core software used for quantum computing simulations is the Qiskit Python library.[5] The Jordan-Wigner [75] encoding method has been considered in each simulation to map the fermionic system on a qubit register.

In table 5.1 we list the data corresponding to the molecules under investigation.

5.2.2 Construction of the Ansatz

In section 5.1.1 we have illustrated how we can define parent sequences of CNOTs to be implemented as an ansatz on the quantum circuit. Anyway, given the arbitrary choice of CNOTs ordering in such sequences, we still have a factorial scaling if we would like to explore all possible orderings. To explore the effects of different choices, for short parent sequences, we have systematically explored all combinations; on the contrary, for parent sequences longer than 6 CNOTs the analysis has been taken on a restricted random selection of ordering. For the randomly chosen ansätze we preferred a uniform distribution for the qubits to be the target of a CNOT. Given the convention in the field of preferring top-down entanglement, the qubits we are referring range from the second to the last one. We proceed as following: we randomly extract t from the set $\{1, \ldots, n\}$ and c from the

set $\{0, \ldots, p-1\}$, then we add the CNOT with control on qubit c and target on qubit t to the entangling block repeating the procedure until it contain $n-1$ CNOTs.
After every application of the chosen entangling block, we apply a $R_Y(\theta)$ quantum gate on every qubit. These two sections of the circuit together are referred to as *block* and with the term *depth* we refer to the number of blocks that will be repeated to build the final ansatz. It is worth noticing that the variational parameters are independent for each block.

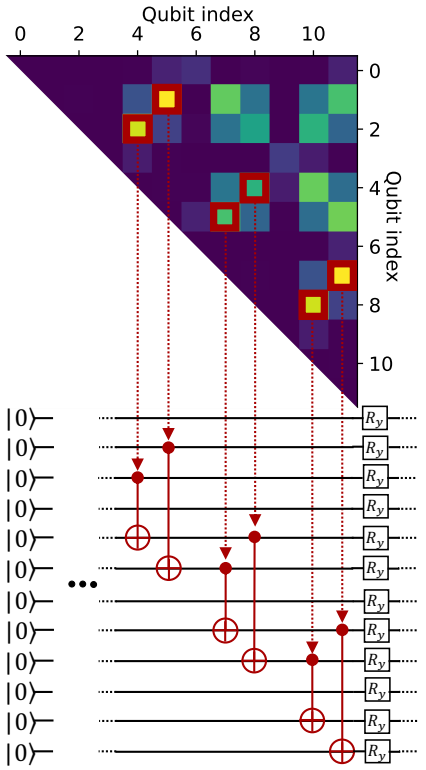


Figure 5.5: Pictorial representation of the method here considered for H_2O molecule. The empirical reduction has been applied to the parent sequence obtained by $\mu = 0.5$. The simple permutation represented in the picture is chosen purely for illustrative purposes. The section of the circuit represented here is then repeated as many times as expressed by the *depth* parameter in the main text.

In the following we show the advantages of our QMI-aided ansatz over the ladder-entangler ansatz, by analyzing the results at different depths. Given the derivation of our algorithm, the resulting ansatz is not generally fully connected, this strengthens its role of starting ansatz, since it can recover a good part of the correlation energy while using a small number of qubits and only after being passed as an overall starting ansatz. The drawback is of course that, with few exceptions, the more expensive ladder-entangler block that links all the qubits will generally give the possibility to span the whole Hilbert space at the price of increasing the gate count.

Table 5.1 shows the values of the μ parameters investigated and the corresponding identified parent sequences. Concerning the choice of the μ parameter, since the systems under analysis are not yet big enough to show a clear lowering of the number of CNOTs as a function of μ, we only consider one value for each system. An exception is considered for the H_2O molecule, for which we used the two values $\mu = 0.5$ and $\mu = 0.7$. The overall seeming unimportance of the parameter μ is justified by the small size of the systems under analysis, indicating that the $\frac{n(n-1)}{2}$ mutual information spots are still too few to extract any long-term general behavior. Furthermore, as shown by Ratini *et al.*[122], theNOfurther contribute to discretizing the distribution of these spots. In general, we expect the importance of the parameter μ to grow with the size of the system, as the sorted arrangement of the $\frac{n(n-1)}{2}$ spots within the $[0, 1]$ range will tend towards continuity.

An example of ansatz construction for the H_2O molecule is given in figure 5.5. It has been obtained by choosing one of the permutations of the parent sequence generated by setting $\mu = 0.5$ and taking into account the empirical reduction. For all the following plots we performed an analysis on 100 different possible permutations of the original parent sequence. Even for the random maps, we performed analysis on 100 of them. In all cases, we performed 50 VQE runs for each entangling block with random starting parameters to guarantee sufficient statistical value for each ansätze. The only exception is the ladder-entangler block, for which we performed 300 VQE since it represents the reference entangling block for comparison. Given its relevance for the results discussed below, the depth 4 of the ladder-entangler ansatz of the water molecule has received 2000 repetitions.

5.3 Results

In this section, we present the results obtained from applying the QIDA strategy to the molecular systems listed in Table 5.1. First of all, we notice that the use of the NO is crucial to increase the sparsity of the I matrix, i.e. to reduce as much as possible the number of qubits pairs with significant values of the QMI. This concept is illustrated in figure 5.1: for each molecule under investigation, we have two maps corresponding to the FCI state expressed in the HFCO and NO basis, respectively on the left and on the central columns. It is clear that the use of the NO is optimal to be used in combination with empirical ansätze since it increases the sparsity of the QMI matrix, reducing the circuit depth, as already pointed out by our previous work [122]. By comparing the images in the central column with those in the right column, one notes the close similarity of the QMI matrices obtained from the FCI and the MP2 states expressed in NO(a similar behavior is also claimed for low bond dimension DMRG calculations [89, 125]), giving consistency to the choice of MP2 state as for input orbitals and QIDA ansätze on the quantum circuit.

Regarding the impact of the different orderings of the parent sequences, we analyze their relative performance in figure 5.6, where we report the results for the H_2O molecule. By setting $\mu = 0.5$, we generated 100 different entangling block orderings from the parent sequence. Each one of them is used as a different ansatz with depth 2, as reported on the horizontal axis of figure 5.6. For each ansatz, we performed 50 independent VQE runs and reported, on the left vertical axis, the maximal value of the percentage of the correlation energy achieved. The right vertical axis gives the probability that the energy of one of the 50 VQEs lies within 30% of the maximum, the value shown on the left vertical axis. Please note that the ordering on the basis of vertical left axes values is not meaningful of any intrinsic property, it was only imposed to facilitate the visual pairing between black and red spots. The data clearly shows how, as expected, some entangling maps perform statistically better than others.

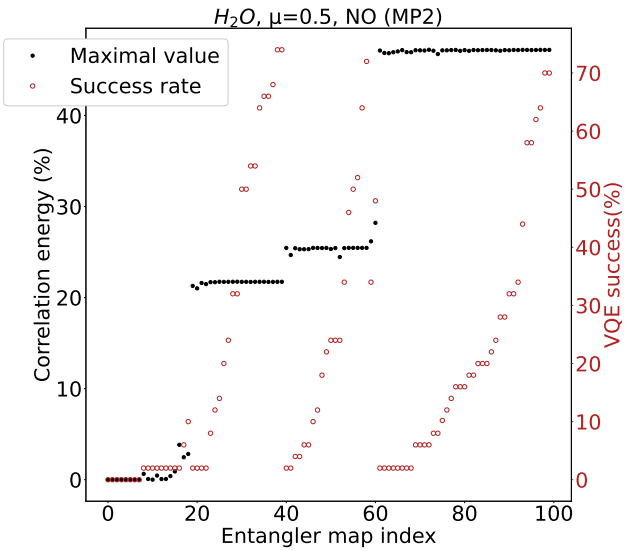


Figure 5.6: Multiple VQEs experiments for the H_2O molecule with different ordering of the same parent sequence corresponding to $\mu = 0.5$. Depth is equal to 2 for each ansatz generated by an entangler map. 50 VQEs runs are performed for each ansatz and the left vertical axis shows the maximum value of the percentage of the correlation energy while the right axes (red) represent the percentage of the 50 VQEs that are within 30% of their respective maximum (black spot). We point out that the rising order in the representation has no objective meaning other than to facilitate the reader to identify the red-black dots pair.

Figure 5.7 is obtained by taking multiple statistical datasets such as the one in figure 5.6 and transposing them over a single column, with this procedure being repeated for all the depths and ansätze shown. Consequently, the width of each column at a given y position represents the probability distribution of the VQE results, similar to a vertical continuous barplot. In each one of these

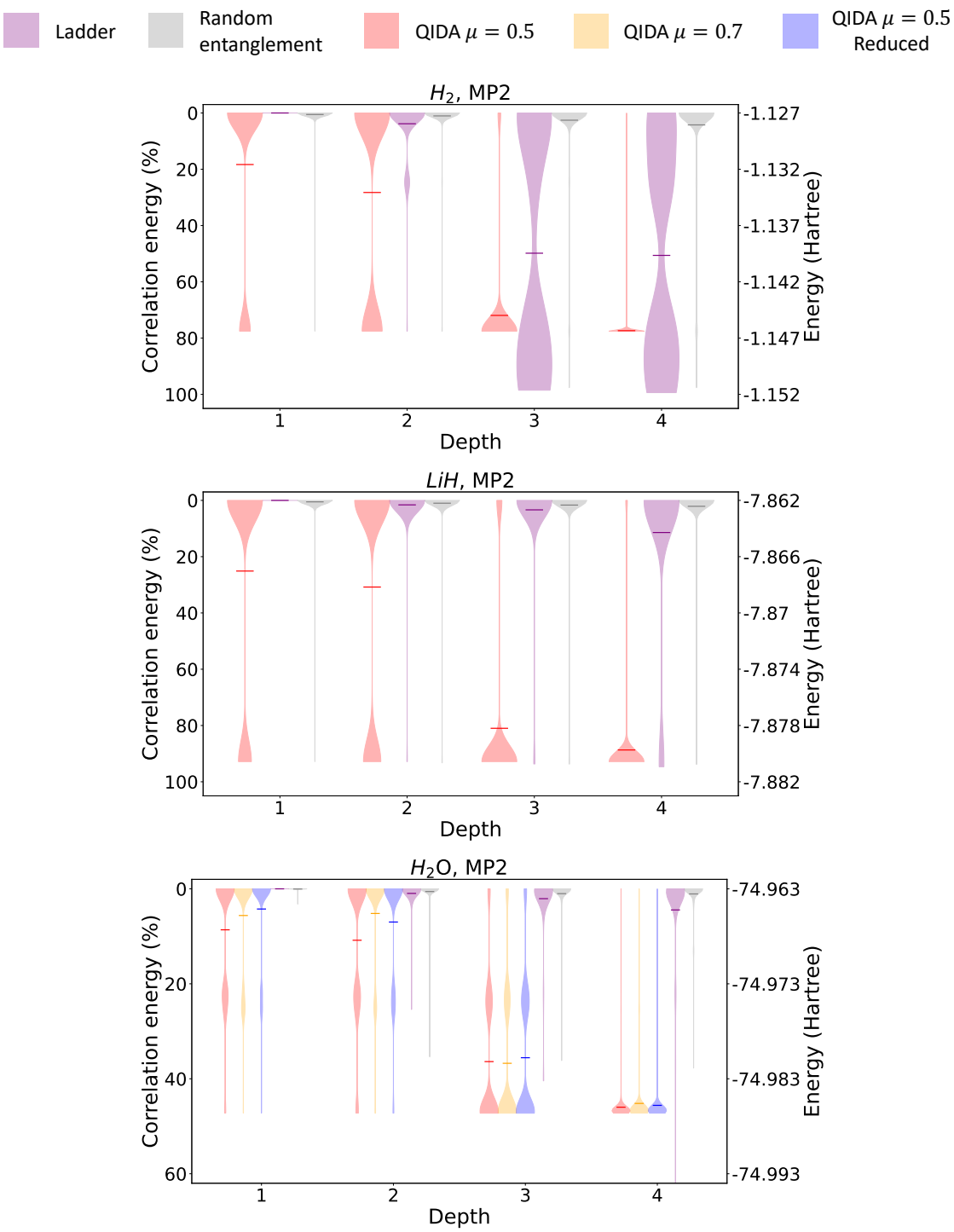
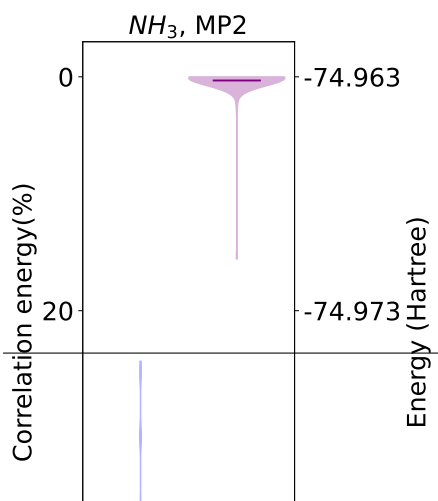


Figure 5.7: Performance of our ansätze compared to the ladder-entangler and random entangler ones. Each vertical color unit (violin) corresponds to the statistical distribution of the energy results for the specified ansatz, with the horizontal dash representing the average. The distributions for QIDA ansätze were obtained by executing 50 VQE runs for each of the 100 permutations of the parent sequence. For the ladder-entangler ansätze, the VQE algorithm has been repeated 300 times, except for the H_2O system with depth 4 (repeated 2000 times). The violins are all normalized by the mode, and only between themselves.

plots, the x-axis defines the depth under analysis, the molecule is pointed out in the title together with the MP2 method from which the orbitals and mutual information is retrieved, the kind of ansätze is then specified by their respective color. For example, the QIDA ansatz at depth 1 for the H_2 molecule is represented in the left-most (pink) column of the first image of Figure 5.1, showing how already at depth one a sizeable percentage of the VQE simulations obtains around 75% of the correlation energy, whereas the average (horizontal line in the column) is still low.

In this way, we effectively estimate the probability distribution (y-axis) of the results obtained by using the QIDA method. Figure 5.7 gives the opportunity to statistically characterize the quality of the results obtained from different parent sequences. For H_2, LiH, H_2O molecules, we considered one or more values for the μ parameter that gives rise to different parent sequences, as reported in table 5.1, which also illustrates if the reduction explained in subsection 5.1.1 has been taken into account. For each parent sequence, we extract 100 random entangling block orderings, as described in the subsection 5.2.2. These blocks are repeated to define ansätze with depths ranging from 1 to 4 and the corresponding VQE results are reported by the violin plots referred to as QIDA in figure 5.7. In such plots the width of the violin is proportional to the occurrences of the energies in the VQE statistics. The different kinds of QIDA ansätze are compared with the violins obtained from the random ansätze (details in subsection 5.2.2) and the ladder-entangler one over depths ranging from 1 to 4. Analyzing the dependence of the maximum value of the percentage of correlation energy on the circuit depth, we see that there is a plateau. Furthermore, already from depth 1 the results spread within the whole interval of the reachable values, with an average value of the correlation energy greater than the one of the ladder-entangler ansatz at depth 4.

Furthermore, we note that the maximal value of the percentage of the correlation energy is obtained by the ladder-entangler ansatz for depths equal to or larger than 3, for the H_2 molecule, and 4, for the LiH and H_2O molecules. By comparing the mean values and the dispersion of the results around it, we see that QIDA ansätze assure an increased probability to obtain better approximations of the FCI state. The rationale behind these results, and thus the explanation for the success of our method, lies in the information encoded in the entangled map based on MP2 calculations. As the ladder-entangler ansatz does not include information about the specific problem for which it is being used, nor is based on any chemically motivated background, the resulting ansatz for the molecular system under investigation, will not necessarily be aimed at finding the optimal set of determinants of the ground state. In the proposed approach here, on the contrary, the topology of correlation is suggested by the precomputed approximated ground state through the QMI. This information directs the optimization towards the space with the expected correlation between qubits, facilitating the optimization and increasing the percentages of success in the VQE runs.



In this regard, for the H_2O molecule, the only system for which we analyzed different values of μ, we did not observe any evidences of advantage in decreasing the value of μ from 0.7 to 0.5, besides a moderate decrease in the expected correlation energy value. This behavior can be explained considering that passing from $\mu = 0.7$ to $\mu = 0.5$ only the three CNOTs between qubits pair 2-8, 2-10, 4-8 are added and, following what has been explained in 5.1.1, these CNOTs can be consid-

ered redundant by the cross-entanglement analysis of the H_2O molecule (fig 5.1f).

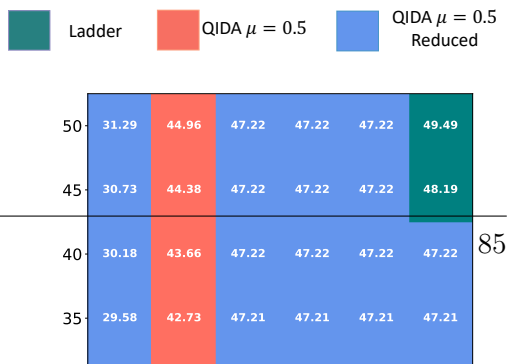
The presented results can be compared with other approaches. Since our goal is to specifically design compact ansätze, we focus our comparison with the field of purely heuristic ansätze, as the approach used by Tkachenko et al. [145], not considering significangly longer, albeit often more accurate, circuits such as those use by Adapt-VQE [64][159]. For instance, in the case of H_2 our QIDA ansatz with depth 1 (MP2) (composed of only 6 CNOTs and 12 parametrized gates) is reaching the value of $\%E_c = 77.5$ with respect to the 8-depth circuit, reaching $\%E_c = 61,4$ (composed of 28 CNOTs and 64 parametrized Ry quantum gates) reported by Tkachenko et al. [145]. Better statistical properties for the convergence of our QIDA ansatz requires a depth of 3 (composed of 18 CNOTs and 20 parametrized gates), which is still arguably a preferable choice. For the case of LiH molecule, we note that percentage of correlation energy reported by Tkachenko et al. [145] using a circuit with depth 8 (composed of 36 CNOTs and 82 Ry) is comparable with our QIDA ansatz since they reach approximately 92% and 94% of the total E_c, respectively. Nevertheless, in the case of QIDA ansatz, the corresponding depth lies between 1 and 3, depending on the statistical distribution desired for the results (as shown in figure 5.7), with a number of CNOTs ranging from 6 to 18 and a number of Ry quantum gates ranging from 14 to 22. In both cases, it is clear how our approach is able to provide a more effective balance between circuit depth and quality of the obtained results, even when a better statistical distribution is required.

Every result shown up to now was simply to give a statistical analysis as extensive as possible to describe the effectiveness of the method described in this work. Having done that, we can now infer what behaviour to expect from other larger systems.

In figure 5.8 we reported the energy results obtained by applying the QIDA method (scheme in figure 5.2) to the NH_3 molecule to compare the outcomes of the VQEs of our scheme with the ones obtained by the ladder-entangler map at depth 4. Thus, for this specific case, we tested our method by repeating 10 VQE runs for 18 different permutations of the parent sequence, while for the ladder entangling maps we performed 200 VQE runs.

Figure 5.8 shows an indisputable advantage for both the maximal value and statistical distribution of the converged VQEs energies. Furthermore, given the CNOTs connectivity of the ansatz, we achieve this result with a reduction in the number of CNOTs from 52 of the ladder-entangler ansatz to 36 of our method. Finally, the trend shown by the molecules in figure 5.7 suggests that the distribution of the reduced QIDA ansatz in figure 5.8 has already reached a plateau corresponding to the maximal value of the correlation energy percentage. Therefore, the ladder-entangler ansatz is still far from reliably finding even a moderate fraction of this quantity.

The previously discussed reduction of the number of needed CNOTs is also relevant for figure 5.7. For example, looking at the bottom figure of 5.7, the reduced QIDA ansatz at depth 2 and the ladder-entangler ansatz at depth 4 have comparable statistical distributions of the energy



results, with the QIDA ansatz having only 12 CNOTs versus 44 CNOTs of the ladder ansatz. Moreover, these distributions have been obtained by repeating the VQE experiments many times. In terms of computational cost, the CNOTs and the number of VQE experiments can be seen as quantum resources, respectively x and y, and we want to investigate the relation between the energy results and these variables. Given a certain parent sequence σ, we introduce the quantity $\epsilon_{\sigma,\delta,i}$ that is the energy obtained at the $i-th$ VQE experiment by the ansatz composed of δ repetitions of σ, the sequence under investigation. Moreover, we introduce the function

$$f_\sigma(x) = max(\delta) \quad s.t. \quad CNOT(\sigma) * \delta \leq x \tag{5.3.1}$$

which, fixed a sequence, finds the maximal number of repetitions that give rise to an ansatz such that the number of CNOTs is less or equal to x. Here, $CNOT(\sigma)$ gives the number of CNOTs corresponding to the sequence under consideration. Finally, for each sequence σ, we define the function

$$\begin{aligned} &E_\sigma(x,y) = \\ &= \left\langle max_{k=1}^{y} max_{\delta=1}^{f_\sigma(x)}(\epsilon_{\sigma,\delta,k}) \middle| max_{k=1}^{y} max_{\delta=1}^{f_\sigma(x)}(\epsilon_{\sigma,\delta,k}) \right\rangle \end{aligned} \tag{5.3.2}$$

that gives the expectation value of the maximal percentage of the correlation energy for each value of the computational resources (x, y) introduced above.

As expected, for each pair (x, y) the function in Eq. (5.3.2) assumes different values by varying σ and the maximal value is reported in figure 5.9. The associated spot has the colour of the corresponding σ, defined in figure 5.7. The advantage of QIDA is remarkable in the overall results with respect to the ladder-entangler ansatz, particularly when the reduction is taken into account. When the number of CNOTs and VQE repetitions is increased, the ladder-entangler ansatz could give better maximal results, without yet increasing significantly the ones obtained with our method. This is not surprising because, as mentioned above, the former connects all the qubits and depends on more parameters, giving the possibility to catch part of the electronic correlations that the latter cannot reproduce even if the depth of the circuit is increased. Overall, statevector simulations give the edge for the fully entangled ansatz only after many repetitions, however, if the noise of a real device was to be taken into account, we assert that the requirement of this ansatz for almost double the count of CNOTs with respect to the QIDA ansatz will likely degrade most of the advantage owned by it in the noiseless case. These positive results encourage the idea of the QIDA ansatz as a modular starting circuit for state preparation that can be followed by a more complete fully entangled circuit that will be able to catch the missing correlations. This is particularly true given that multiple repetitions of the QIDA ansatz seem to guarantee the best reachable energy

for a limited starting connectivity, regardless of the chosen permutation, thus allowing for a good starting point for successive rounds of VQE. On the other hand, if the requirement concerns the shortest possible circuit at the price of worse energies, QIDA is still the best choice, given that it allows for good outcomes even at depth 1 (as seen in figure 5.7).

5.4 Conclusions

In this paper, we have presented a novel approach to constructing hardware-efficient ansätze for molecular systems. Our method uses quantum mutual information from the wavefunction calculated by a well-scaling classical method, like MP2, to create initial ansätze for variational quantum algorithms that achieves correlation energies comparable or superior to currently used ladder-type heuristic ansätze, while requiring substantially fewer CNOTs.

To demonstrate the effectiveness of our approach, we studied various molecular systems using statevector simulations. For the H_2O molecule, with just 24 CNOTs, our QIDA method achieved an expectation value of the correlation energy of one VQE of 45.56%, significantly outperforming the ladder-entangler ansatz, requiring 44 CNOTs and yielding only 4.5% of the correlation energy on average. It is worth stressing that the results in figure 5.9 have been obtained with statevector simulations. Then, for noisy simulations, we expect that QIDA ansätze will become even more efficient with respect to the ladder-entangler one because the number of two-qubits gates is significantly reduced. Additionally, we observed a growing performance disparity with the number of qubits (see fig. 5.8), with our QIDA method consistently demonstrating superior performances compared to the ladder-entangler ansatz.

Furthermore, we propose a modular approach to use our QIDA method in conjunction with other ansätze. The idea is to first employ QIDA to identify the leading correlations between qubits in order to deliver optimized ansätze with essential albeit appropriate starting ingredients, excluding entanglement between loosely correlated qubits. The state obtained running the VQE algorithm with such ansatz can be used as a starting point for subsequent application of entangling blocks that may refine the ansatz by catching the remaining correlation. This approach has the potential to reduce the overall computational cost of quantum chemistry simulations while maintaining or improving the accuracy of the results.

In summary, our work has introduced a promising method for generating hardware-efficient ansätze for quantum chemistry inspired by quantum information. We have demonstrated the effectiveness of our QIDA method on a set of small molecules, increasing the percentage of the correlation energy values collected with fewer CNOTs compared to existing ansatz methods.

Chapter 6

Compact Multi-Threshold Quantum Information Driven Ansatz For Strongly Interactive Lattice Spin Models

In the previous chapter, we introduced a new kind of Hardware efficient Ansatz, based on approximated knowledge of the state of interest. In this chapter, we further develop this idea by applying it to different variations of the Heisenberg model and by reproducing this idea over different iterations for the development of more refined and adaptable Ansätze. Based on the work Ref. [143]

Considering the previous discussion of 5 for improving the VQE, various algorithms have been developed [14, 18, 64, 156] and are present in the literature. Specifically, considering the focus of the field on designing effective ansätze , it comes as natural that analogous designs specific for strongly correlated lattice model, like the Heisenberg Hamiltonian, have been developed. An example is the Hamiltonian Variational Ansatz (HVA) which leverages the compact structure of lattice model Hamiltonians, using fermionic operators to exploit low Pauli weight encodings [38, 67, 152]. This approach often results in shallower circuits compared to general-purpose ansatze.

In contrast, the second approach, the Heuristic (or Hardware efficient) Ansatz[62, 76, 124, 142, 145], is designed without relying on information about the physical system, focusing solely on exploiting the quantum hardware's capabilities. As already explained in this thesis, while the Heuristic Ansatz better utilizes quantum hardware, it comes at the cost of losing the physical meaning associated with the variational ansatz.

Independently from the criterion according which the ansatz is built, as the Parameterized Quantum Circuit (PQC) start to get increasingly complex, the length and depth of the circuits increase too, not only enhancing the probability of accumulating errors but also leading to the phenomenon of Barren Plateau [86] i.e. the exponentially flat parameter optimization landscape. While modification in the optimization methods or additional error mitigation are not strategies to avoid barren plateaus, shallow and variable structured ansätzes are efficient in the reduction of the phenomenon. Examples of these solution are ADAPT-VQE ansätze [64], which grows iteratively the complexity of the guess wavefunction one operator at time following information obtained by the simulated molecule. Exploiting a combined iterative optimization of the Hamiltonian and the wavefunction, the WAHTOR algorithm shows the capability of reducing the depth of the circuits and reduce the risk of falling into local minima while reaching a lower variational energy [120]. Similarly

to the adaptive approach, the Quantum Information Driven Ansatz (QIDA) [96] introduced in the previous chapter 5 provides a quick and fast way to generate an initial guess with few CNOTs.

The aim this work is to extend further the idea behind the QIDA approach and to expand the field of application beyond quantum chemistry. The main focus of the QIDA method was to utilize Quantum Mutual Information (QMI) to define shallow-depth circuits. QMI, traditionally used in quantum chemistry to measure the correlation between orbitals [48], has found applications in quantum computing such as optimizing qubit ordering [145] and enhancing quantum algorithms [159]. The circuits obtained by the QIDA approach represent compact and affordable initial starting states from which more precise and complex ansätze can be built.

Following the QIDA approach, our state-preparation method aims to recover the missing correlation by exploiting a layered-structured ansätze, where each layer is selected based on the QMI values. Multi-Threshold Quantum Information Driven Ansatz (Multi-QIDA) extends the single threshold QIDA approach by including qubit-pairs that still present mid-high QMI values, that are excluded by the original method. Combined with a repeated iterative VQE routine, Multi-QIDA allows to obtain effective shallow circuits that surpass the generic ladder-fashion ansatz in terms of performance.

This work mainly develops on the previous QIDA algorithm of 5 to improve its performance with a focus on the Heisenberg Model Hamiltonian, introduced in Sec. 6.1.1. In Sec.6.1.2 we then introduce the $\mathbf{SO}(4)$ gates used as alternatives to the standard CNOTs ansatz. Our complete approach, Multi-QIDA, is introduced in Sec.6.2. The procedure to convert quantum mutual-information matrices into a sequence of entangling gates is defined in Sec.6.2.1. Then, the construction of the parametrized quantum circuit representing the variational form introduced by our approach is explained in Sec.6.2.2. The iterative optimization procedure combined with the multi-layer ansatz is defined in Sec.6.2.3. Simulation details are presented in Sec.6.3. Finally, in Sec.6.4 results are shown and explained.

6.1 Method

6.1.1 Heisenberg Model

Our method's efficiency was evaluated through benchmarking on various configurations of the Heisenberg model Hamiltonian. The systems selected for this analysis encompassed diverse settings, providing a comprehensive assessment of our approach across different scenarios.

The most general way to define the Heisenberg model Hamiltonian is the following:

$$H = \sum_{\langle i,j \rangle} \mathbf{J}_{i,j} \mathbf{S}_i \cdot \mathbf{S}_j + \mathbf{h}_i \sum_{i}^{N} \mathbf{S}_i, \tag{6.1.1}$$

where $\langle i, j \rangle$ indicates all nearest-neighbor spins, $\mathbf{J}_{ij}$ is the interaction term between two adjacent spin sites, $\mathbf{S}_i$ and $\mathbf{S}_j$ are spins operators acting on sites i and j, respectively, and $\mathbf{h}_i$ is the local magnetic field acting on the i-th site and this can be an external magnetic field or magnetic impurities. Expanding each spin operator $\mathbf{S}_i$ into the three components along x, y and z axis, the Hamiltonian

H can be written as

$$H = \sum_{\langle i,j \rangle} (J_{ij}^{xx} S_i^x S_j^x + J_{ij}^{yy} S_i^y S_j^y + J_{ij}^{zz} S_i^z S_j^z) + h \sum_i^N S_i^z, \tag{6.1.2}$$

where $\langle i, j \rangle$ indicates all nearest-neighbor spins in a 2D lattice with open boundary conditions, $J_{ij}^{\beta\beta}$ is the interaction term between two adjacent spin sites with $\beta \in \{x, y, z\}$, h is the external magnetic field and S_i^β is the β component of the $\frac{1}{2}$-Spin operator acting on the i-th spin site.

The Hamiltonian is written directly on the qubit space by converting each $\frac{1}{2}$-Spin operator into Pauli operator following

$$S_i^\beta = \frac{\hbar}{2} \sigma_i^\beta, \tag{6.1.3}$$

where $\sigma_i^\beta \in \{\sigma^x, \sigma^y, \sigma^z\}$ i.e one of the three Pauli matrices. The Hamiltonian to be measured on a quantum computer must be converted into a sum of Pauli strings. A Pauli string is defined as a tensor product of Pauli operators. Each Pauli string corresponds to one of the coupling terms or the on site term in Equation 6.1.2. For a general coupling term, $J_{i,j}^{\beta\beta}$, the relative Pauli string is defined as

$$I_0 I_1 \dots I_{i-1} \sigma_i^\beta I_{i+1} \dots I_{j-1} \sigma_j^\beta I_{j+1} \dots I_{N-1}, \tag{6.1.4}$$

where I_i are identity matrices.

6.1.2 SO(4) Gates

Our approach has been also tested with a more structured correlator, instead of only CNOTs and R_y rotations, a fully parametrized from the $\mathbf{O}(4)$ group. In particular, we have chosen to implement two-qubit gates with determinant equals 1, which corresponds to two-qubit gates from $\mathbf{SO}(4)$.

Differently form standardize CNOTs based ansatz, in which the correlators can be seen as a switch, you insert them or you do not, the parametrized $\mathbf{SO}(4)$ gates offer a tunable correlation, that can be chosen by the optimization. In a sense, they offer partial rotations and intermediate transformations that are not available whose are not directly available with CNOTs only. With the right parametrization, the group of $\mathbf{SO}(4)$ transformations can also lead to the identity, and also symmetry preserving gate, e.g. the fermionic swap Eqn.6.1.5, which allows to impose anti-symmetrization to the ansatz.

$$fSWAP = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & -1 \end{pmatrix} \tag{6.1.5}$$

On top of the variety of intermediate transformation that can be obtained by parametrized $\mathbf{SO}(4)$ matices, the decision of this type of gates is also related to the nature of the system that we decided to treat. Given the time reversability of the Heisenberg model Hamiltonians thus we only need to include real transformations described by matrices in the $\mathbf{O}(4)$ group and $\mathbf{SO}(4)$. In the context of $\mathbf{O}(4)$ and $\mathbf{SO}(4)$, the matrices consist of real numbers, and the operations involved in this group are based on real arithmetic, both can represent rotations in four-dimensional space.

Following Theorem 3 in [146], every two-qubit quantum gates in $\mathbf{SO}(4)$ can be realized by a circuit consisting of 12 elementary one-qubit gates and two CNOT gates.

A generic gate $U \in \mathbf{SO}(4)$ is composed by two generic one-qubit rotations $A, B \in \mathbf{SU}(2)$, four S gates and two R gates. It is also known that every matrix $M \in \mathbf{SU}(2)$ can be written as a composition of $R_z(\alpha)R_y(\theta)R_z(\beta)$ for some α, β and θ, while the R gate is defined as $R_y(\pi/2)$ and the S gate is obtained with $R_z(\pi/2)$. The U gate is then parametrized using the two set of three parameters of gates A and B.

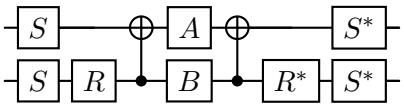


Figure 6.1: Circuit implementing a general $U \in \mathbf{SO}(4)$.

6.2 Multi-QIDA

The main focus of the Multi-QIDA approach is the selection of the missing correlation that is not caught when applying the QIDA method of [96] as a standalone ansatz, which has been explained in detail in Sec. 5.

The implementation of Multi-Threshold Quantum Information Driven Anstaz (Multi-QIDA) is combination of three steps, which are described in the following sections. The initial step is the *Layers-Builder* procedure, a purely classical routine which, given a QMI matrix, infers the adaptive layers of the VQE by sectioning the MI matrix by correlation levels. Second, the *Circuit-composition*, the circuit made of the different entangler maps for each layer is compiled, with details varying based on the entangler block chosen, $CNOT$ or $\mathbf{SO}(4)$. Finally, we perform an *iterative layered-VQE*, a varied version of the Layer-VQE [91] which converges to the full ansatz.

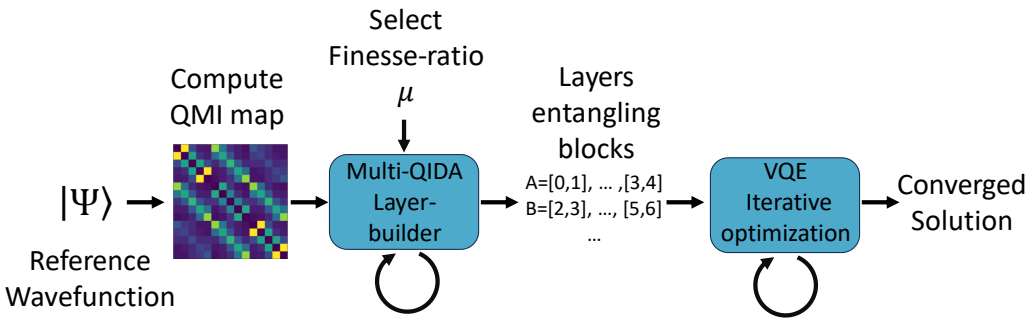


Figure 6.2: Pictorial representation of the workflows of the proposed Multi-QIDA algorithm.

6.2.1 Layers-Builder

To improve over the QIDA method, which mainly takes into account low correlated qubits pair, we iteratively exploit QMI maps to obtain other additional layers. The selection is performed on the list of qubits-pairs ordered for decreasing QMI value from the QMI matrix $I_{i,j}$. Using a threshold $\mu \in [1, 0)$, chunks of qubits-pairs that a layer must contain can be selected. This threshold μ, called *finesse-ratio*, represents the range of QMI values that are desired to be inserted in each additional layer.

Algorithm 2: Schematic outline of the Multi-QIDA Layers-builder

Input : QMI matrix I_{ij}, $\bar{\mu}$, $N_{\text{qubits}} > 0$
Output: List of entangling map L

$L \leftarrow$ empty list;
$free \leftarrow \{n \mid \forall n \in N_{\text{qubits}}\}$;
$G \leftarrow$ Disconnected_Graph($free$);
$t \leftarrow 1$;
$all_connected \leftarrow$ false;
while $(free \neq \emptyset)$ **and** $\neg(all_connected)$ **do**
    $l \leftarrow$ empty list;
    for $q_i, q_j \in \{V_{I_{i,j}} : \bar{\mu}[m] > I_{i,j} > \bar{\mu}[m+1]\}$ **do**
        if $\neg(G.connected(q_i, q_j))$ **then**
            l.add($[q_i, q_j]$);
            G.add_edge(q_i, q_j);
            $free$.remove(q_i, q_j);
        end
    end
    L.append(l);
    $t \leftarrow t - \mu$;
    $all_connected \leftarrow G$.is_connected();
end
L.append(ladder_ansatz(N_{qubits}));
return L

Initially, all the qubits are in the set of *free qubits*, meaning they have not been used in any entangling process. A QIDA entangling layer l-th is then composed by selecting the pairs q_0, q_1 whose correlation I_{q_0,q_1} is in the interval $[1 - l\mu, 1 - (l+1)\mu)$. Such pairs are then added to the candidate entangling list of the l-th layer, which represents the potential pairs to which then impose a correlator. The criterion according which the entangling gates are added to the current layer l is defined as follow: If the two qubit q_0 and q_1 are not reachable each other through direct connection or exploiting previous layers qubit-pairs, they are selected. Thus, they belong to distinct subsets of qubits. As a consequence, if at least one of the qubits was *free* up to the current layer l, then this is clearly automatically satisfied. This procedure is iterated again until the set of free qubits is empty and all the qubits belong to the same unique subset, i.e. they are all connected. In other terms, the layers are added until all the qubits in the lattice have been covered by at least one entangling layer. After the last layer is defined with this procedure, an additional layer of entanglers in the ladder setup closes the Multi-QIDA ansatz.

At the end of this building procedure, we define the entire collection of layers by L, which is then passed to the ansatz construction procedure.

6.2.2 Circuit ansatz composition

Until now, we only took care of retrieving how many correlators to impose on the quantum circuit and where to impose them, now, starting from the list of entangling maps L, we construct the ansatz in two ways. The first implementation follows the standard combination of parametrized single-qubit rotation gates and $CNOT$ as entanglers [76], and it has been included in the analysis as a

comparison to the next type of ansatz. In the following Multi-Qida construction with this entangler will be referred to as $QIDA^{CX}$. The second kind of ansätze is obtained by entangling pairs of qubits using $\mathbf{SO}(4)$ gates, as in Fig. 6.1, which, as we have seen, guarantees the most general description of two-qubit gates, such Multi-Qida implementation are referred in the following as $QIDA^{SO4}$.

As already introduced in the first section, the phenomenon on the barren plateau is related to the *circuit expressiveness*: the larger is the variational space that the ansatz has to explore, the greater is the probability, when initialized according to a random distribution of the parameters, to lead to initial states that are far from the right solutions. One of the possible solutions to avoid Barren Plateau is to employ short PQC [86]. By variationally optimizing only a portion of the search space with the restricted parameters, in some scenarios, shallow circuits present no flattened potential surfaces, with the drawback that the optimal solution could also not be obtained inside this constrained space. Following this concept, similarly to the Adapt-VQE [64] we decided to extend the shallow ansatz obtained by the QIDA method, adding layer-by-layer expressiveness to the PQC, without exploring the huge variational space at once. At each $l \neq 1$, we want the new layer L_l to beinitialization to the identity in order to not disrupt all the previous $l-1$ rounds of optimizations once a new (yet) unoptimized layer is introduced.

Specifically, for the combination of R_y and CNOT, we can obtain the identity by imposing a V-shape structure to the entangling layer l, partly resembling the topology of a UCC excitation [40, 53, 68, 94]. The V-shape is obtained by a series of R_y rotations, the additional entangling layer, a central layer of R_y rotations, again the additional entangling layer but with reversed-order, completed with a last series of R_y rotations, as shown in Figure 6.3.

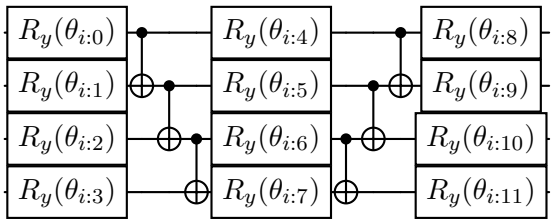


Figure 6.3: Circuit implementing a V-shape i-th additional layer with entangling map $[0,1]$,$[1,2]$,$[2,3]$. This circuit results into an identity, $\mathbb{1}$, when all the parameters $\theta_{i:j}$ are zero for all j.

If the additional layer is built with $\mathbf{SO}(4)$, only $\mathbf{SO}(4)$ gates are used in the newly added circuit , and since the identity gate is already included in the zero-valued parametrization of a general $\mathbf{SO}(4)$(Fig. 6.1), no further construction is needed to obtain it in this case. We then notice that with both kinds of entangling map (Fig. 6.4 and Fig. 6.3), setting the parameters of to zero guarantees an identity gate.

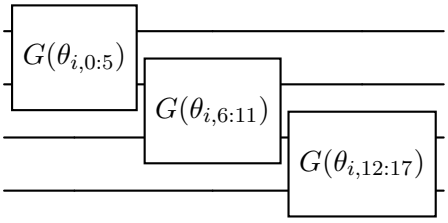


Figure 6.4: Additional layer with entangling map $[0,1]$,$[1,2]$,$[2,3]$ where each $G \in \mathbf{SO}(4)$. This layer results into an identity, $\mathbb{1}$, when all the parameters $\theta_{i,j}$ are zero for all j.

The initialization of the parameters of the additional layers with zeros, or in general with a circuit that results in an identity, ensures that the cost function being optimized remains unchanged after introducing this new layer. This structure is applied to all the entangling layers contained in L, except for the first one, as in that case there is yet no circuit to disrupt. Thus the first layer does not need to include an identity in its parametrization and as such is defined as a standard heuristic ansatz with random initial parameters.

6.2.3 Iterative layered-VQE

The selection of a suitable optimization routine is necessary when dealing with different circuit layouts in which the composition of the ansatz may vary from layer to layer. As the name of this section suggests, this will be done iteratively along many steps, each of which will include two main phases, an optimization of the single layer L_l at the l-th step, and a global optimization of all the previous layers $\{0\dots,l-1\}$. We remind the reader that each layer L_l can be composed of two different kinds of entangler, as explained in 6.2.2, however, the procedure explained here is independent of this variable. Furthermore, as was specified before, any layer but the first one ($l>0$) needs to be initialized to the identity gate, which implies that the parameters of both $QIDA^{CX}$ and $QIDA^{SO4}$ ansätze are initialized to zero, for the first layer ($l=0$), the parameters are instead initialized randomly in the interval $[0,2\pi)$.

The starting parameters, denoted as $\bar{\theta}_l$, represent the unoptimized parameters associated with the l layer. With the vector of parameters $\bar{\theta}_l$, it is possible to define the unitary transformation to bring $|\Psi_l\rangle$ to the $l+1$-th state $|\Psi_{l+1}\rangle = U(\bar{\theta}_l)\,|\Psi_l\rangle$. To optimize the $\bar{\theta}_l$ parameters we apply the VQE algorithm by minimizing $\langle\Psi_l|\,U^\dagger(\bar{\theta}_l)HU(\bar{\theta}_l)\,|\Psi_l\rangle$. We mark the optimal parameter vector by $\bar{\theta}_l^*$.

Now we may unfold the description of the total state $|\Psi_l\rangle$ by listing all the previously optimized layers with $|\Psi_{l+1}\rangle = U(\bar{\theta}_l^*)\,|\Psi_l\rangle = \prod_{j=0}^{l} U(\bar{\theta}_j^*)\,|\Psi_0\rangle$. After the completion of optimization of the layer L_l the algorithm proceeds with a new VQE. In this second phase of the iteration l the algorithm encompasses all the layers $\{0\dots,l\}$ in a general VQE that minimizes $\langle\Psi_0|\prod_{j=l}^{0} U^\dagger(\bar{\theta}_j)H\prod_{j=0}^{l} U(\bar{\theta}_j)\,\langle\Psi_0|$. This optimization takes as a starting point the previous parameter vectors and gives as a result a new set of different parameter vectors $\bar{\theta}_j^*\forall j\in\{0\dots,l\}$.

Defining a general procedure for the l-th layer in L and exploiting a saved empty circuit QC_{empty} i.e. PQC composed by layers up to the $(l-1)$-th, we append the l-th layer to the QC_{empty}, after assigning the optimal parameters up to the previous iteration $\bar{\theta}_{0,1,\dots,l-1} = \bar{\theta}^*_{prev}$. The l-th layer parameters are set to zero. We find the optimal set of parameters of the l-th layer alone, $\bar{\theta}_l^*$. The total set of parameters up to the l-th layer is composed $\bar{\theta}_{tot} = \bar{\theta}^*_{prev} + \bar{\theta}^*_l$ (where $+$ is not an addition but a concatenation of the two sets). Finally, the QC_{empty}, which contains all the layers up to the l-th, is used as the ansatz for the final VQE where the initial starting parameters is $\bar{\theta}_{tot}$, obtaining the optimized set of the combined circuit i.e. $\bar{\theta}^*_{tot}$.

Algorithm 3: Iterative (Re)-Optimization routine

Input : $N_{\text{qubits}} > 0$, List of entangling map L

Output: Optimal parameters $\bar{\theta}^*$, Converged energy E

$QC_{\text{empty}} \leftarrow \text{QuantumCircuit}(N_{\text{qubits}})$;
for $l \in L$ **do**
  append(QC_{empty}, l);
  $M \leftarrow \text{n_params}(QC_{\text{empty}})$;
  if l *is first layer* **then**
    $\bar{\theta}_0 \leftarrow \theta_{0:0}, \ldots, \theta_{0:M} \in \mathbb{R}\,[0, 2\pi]$;
    $\bar{\theta}^*_l \leftarrow \text{VQE}(QC_{\text{empty}}, \bar{\theta}_0)$;
  else
    $\bar{\theta}_l \leftarrow \theta_{l:0}, \ldots, \theta_{l:M} = 0$;
    $\bar{\theta}^*_l \leftarrow \text{VQE}(QC_{\text{prev}}, \bar{\theta}_l)$;
    $\bar{\theta}_{\text{tot}} \leftarrow \bar{\theta}_{\text{prev}} + \bar{\theta}^*_l$;
    $E_{\text{tot}}, \bar{\theta}^*_{\text{tot}} \leftarrow \text{VQE}(QC_{\text{empty}}, \bar{\theta}_{\text{tot}})$;
    $\bar{\theta}^*_{\text{prev}} \leftarrow \bar{\theta}^*_{\text{tot}}$;
    $QC_{\text{prev}} \leftarrow \text{assign}(QC_{\text{empty}}, \bar{\theta}^*_{\text{prev}})$;
  end
end
return $E_{\text{tot}}, \bar{\theta}^*_{\text{tot}}$

We notice that the optimization procedure shown in Algorithm 17 is not limited to the use in combination with Multi-QIDA, but could also be used as an optimization method for other iterative ansätze, as it is a similar procedure to the one used by the Adapt-VQE algorithm, as well as in Layer-VQE.

6.3 Computational details

6.3.1 Spin Models

To benchmark our experiments, we have used a total of six configurations i.e. variations of the original Heisenberg Hamiltonian in Equation 6.1.2. The different variations can be resumed in the following Hamiltonian:

$$H = \frac{J}{4} \sum_{\langle i,j \rangle} [\Delta(\sigma_i^x \sigma_j^x + \sigma_i^y \sigma_j^y) + \sigma_i^z \sigma_j^z] - \\ - \frac{h}{2} \sum_{i}^{N} \sigma_i^z. \tag{6.3.1}$$

In Equation 6.3.1, the operator is defined in terms of three parameters: $J = J_{ij}^{\beta\beta}$ for all $i, j \in \{1, \ldots, N\}$, $\beta \in \{x, y, z\}$ and $J > 0$, which is the interaction term, Δ as the Anisotropy term, and the external magnetic field h.

- *Isotropic Systems*: the first three experiments were carried on the the *Isotropic anti-ferromagnetic* Heisenberg model Hamiltonian. This Hamiltonian is obtained form Equation 6.3.1, imposing

$J = 1$, $\Delta = 1$, and $h = 0$.

- *Anisotropic Systems*: the system studied is called *XXZ Anisotropic* Heisenberg model Hamiltonian and appears when the coupling terms associated with J^{xx} and J^{yy} are equal but different from J^{zz} i.e. $J^{xx} = J^{yy} \neq J^{zz}$. The anisotropy can also be defined as $J^{xx} = J^{yy} = \Delta J^{zz}$, where $0 < \Delta < 1$, from which is easy to see that with $\Delta = 1$ the isotropic Hamiltonian is obtained and an Ising Hamiltonian setting $\Delta = 0$ can be built. The XXZ Hamiltonian used in our simulation are defined with a $\Delta = \{\frac{1}{10}, \frac{2}{3}\}$, $J = 1$, and $h = 0$.

- *Isotropic with external magnetic field system*: the last configuration tested is an isotropic Heisenberg Hamiltonian with $J = 1$ and external magnetic field $h = 2$.

In terms of sizes, we have used three lattices to test the standard isotropic Heisenberg Hamiltonian, namely 3×3, 2×6 and 3×4. For the variation of the Hamiltonian configuration i.e. isotropic with external magnetic field and anisotropic, we decided to test them only for the 3×4 lattice.

Lattice	*#Qubits*	J	h	Δ	E_{exact}	E_{neel}
3×3	9	1.0	0.0	1.0	-4.749327	-3.00
2×6	12	1.0	0.0	1.0	-6.603472	-4.00
3×4	12	1.0	0.0	1.0	-6.691680	-4.25
3×4	12	1.0	2.0	1.0	-9.508473	-4.25
3×4	12	1.0	0.0	0.66	-5.338751	-4.25
3×4	12	1.0	0.0	0.1	-4.272670	-4.25

Table 6.1: Summary table containing all the configurations used in the experiments. The quantities J, h, and Δ have been defined in Section 6.1.1, E_{exact} is the energy obtained by exact diagonalization of the Hamiltonian of the system. The energy of the reference state, E_{neel}, is obtained following Equation 6.3.4.

6.3.2 Metrics and measure

We employed various metrics and measures to compare performances across different systems. In addition to the straightforward metric of the number of CNOTs, $\#CNOT$, other measures included are Absolute Quantum Energy, Relative Quantum Energy, Minimum Absolute Energy Deviation, and Minimum Relative Energy Deviation. While $\#CNOT$ is an unambiguous definition, other measures need a brief explanation:

- *Absolute Quantum Energy* (AQE_i) for the i-th VQE run, it is calculated as the ratio between the energy obtained from the i-th VQE run and the one obtained from Exact Diagonalization, E_{exact}.
$$AQE_i = \frac{E_i}{E_{exact}} * 100. \tag{6.3.2}$$

- *Relative Quantum Energy* (RQE_i) is defined similarly as the ratio between the difference of the i-th VQE run and the Néel energy and the difference between the ED energy and the Néel energy:
$$RQE_i = \frac{|E_i - E_{neel}|}{|E_{exact} - E_{neel}|} * 100. \tag{6.3.3}$$

The Néel state represents the ground state of the corresponding classical system, and it is obtained by setting spin in an alternated configuration between spin-up, $\uparrow$, and spin-down, $\downarrow$. The

energy associated with the Néel state, E_{neel}, is a constant that scales with the number of pairings and the value of the coupling term J_{zz}, i.e.

$$E_{neel} = \sum_{\langle i,j \rangle} J^{zz}_{i,j}. \tag{6.3.4}$$

All the values relative to E_{exact} and E_{neel} are shown in Table 6.1.

- *Minumum Absolute Energy Deviation* (MAED) represents the average deviation of the absolute quantum energy from the best achieved AQE i.e. AQE_{best}, denoting the AQE value of the top-performing run. It is computed by summing the differences between the AQE of each VQE run and the AQE_{best}, then normalizing this sum by the total number of runs. It is expressed as:
$$MAED = \frac{\sum_{i=1}^{\#VQE} |AQE_i - AQE_{best}|}{\#VQE}. \tag{6.3.5}$$
- *Mininum Relative Energy Deviation* (MRED) quantifies the average deviation of the correlation energy from the optimal energy E_{best}. It's calculated by summing up the differences between the Relative Quantum Errors RQE of each VQE run and the RQE_{best}, which represents the lowest RQE achieved among all runs. This sum is then normalized by the total number of runs.
$$MRED = \frac{\sum_{i=1}^{\#VQE} |RQE_i - RQE_{best}|}{\#VQE}. \tag{6.3.6}$$
- *Minimum Energy Deviation* (MED) defined as the sum of the absolute difference between each VQE run and the best-performing run, E_{best} corresponding to the minimum obtained energy, for the same configuration. The value is then normalized using $\#VQE$, which is the number of experiments.
$$MED = \frac{\sum_{i=1}^{\#VQE} |E_i - E_{best}|}{\#VQE}. \tag{6.3.7}$$

6.3.3 Simulation details

We split the workflow between pre-processing and simulation. For the preprocessing part, which includes the DMRG calculation and QMI creation, we used `ITensors.jl` [58, 59], a supplementary Julia library for efficient tensor computations and tensor network calculations. For the Qubit-pairing selection and the framework for the simulations, we used Python, while for the quantum computing simulation and PQC creation, we used the Qiskit Python library[5]. We tested every Hamiltonian configuration with 50 VQE runs each. We employed noiseless statevector simulation. For the optimization procedure of the VQE, we decided to use the Broyden–Fletcher–Goldfarb–Shanno (BFGS)[60] algorithm with a convergence threshold set to 10^{-6}.

6.4 3x4 Isotropic System Results

In this section, we will present the detailed workflow and results obtained for the 3×4 Isotropic system. We have then collected the complete results for all the Hamiltonian configuration in the Appendix.

6.4.1 QMI Map and Qubit-pairs selection

First of all, we employed DMRG to obtain the reference state i.e. the MPS representing the approximation of the Heisenberg Hamiltonian groundstate. Starting from the DMRG reference MPS, we created the corresponding QMI map. We compute each term of Equation 1.10 by exploiting properties of the MPS wavefunction that allow us to compute traced density matrices matrices, defined in Equation 1.7, efficiently. In Fig.6.5, we show the QMI map obtained for the 3×4 isotropic Heisenberg Hamiltonian.

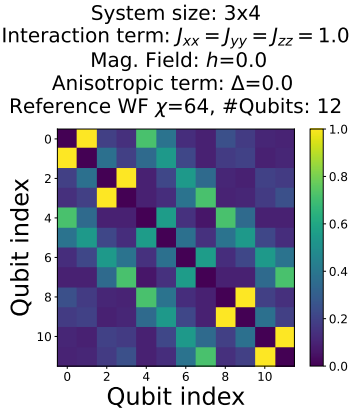


Figure 6.5: 3×4 Isotropic Heisenberg model Hamiltonian QMI map.

From the QMI map, we have extracted the qubit-pairing and ordered them in ascending mutual-information value. Then, we applied the layer-building procedure explained in Section 6.2.1, following Algorithm 20. In Figure 6.6, we show a pictorial representation of the ordered set of coupling obtained by a QMI map. In the upper plot, all the qubit-pairs are represented, and the color of each

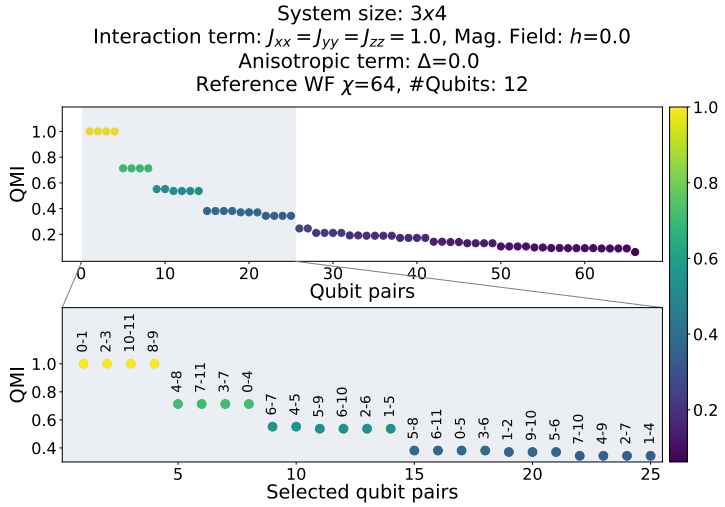


Figure 6.6: 3×4 QMI pairings selection.

point is chosen accordingly with the value of QMI that it possesses. The most correlated couplings are the ones with a QMI value near 1, colored yellow, and a decreasing value of QMI corresponds to a color gradient towards dark blue. In the bottom part of the plot, the selected qubit pairs contained in the darker area are displayed in detail. The selected pairs correspond to the highly correlated pairs accordingly to the QMI values. The selection has been performed with a finesse

ratio $\mu = 0.1$ and the visualization encloses also qubit-pairs that will be removed in the optimized addition of entanglers in the Multi-QIDA layers-builder algorithm. These pairings are the ones that try to connect qubits that are already connected by the previous layer and thus, they are reachable to each other by exploiting the cross-correlation of the previous QIDA layers.

6.4.2 Layer-building procedure

From the set of qubit-pairs split accordingly to the QMI values, we retrieved the actual Multi-QIDA ansatz that we used as PQC. After the selection procedure, we defined an ansatz that contains only a small number of pairings. For this system, we employed only less than one third of the total couplings available. The isotropic 3×4 model Hamiltonian required 17 of the 66 available qubit-pairs. The 17 correlators are split in the following layers :

- First Layer (4 correlators):[0, 1], [2, 3], [8, 9], and [10, 11].
- Second Layer (4 correlators): [0, 4], [3, 7], [4, 8], and [7,11].
- Third Layer (6 correlators): [1, 5], [2, 6], [4, 5], [5, 9], [6, 7], and [6, 10].
- Fourth Layer (3 correlators): [1, 2], [5, 6], and [9, 10]
- Fifth Layer (11 correlators): final ladder.

Following the above composition, we obtain a PQC that is composed of 52 CNOTs for $QIDA^{CX}$ and 56 CNOTs for $QIDA^{SO4}$ in a total of 5 QIDA layers. Assuring that only pairing containing qubits that are not already connected, we exploit cross-correlation between already connected qubits thanks to previous layers, and in this way we have reduced the number of qubits-pairs required in each additional layer. For the isotropic 3×4, we exploited this optimization to reduce the number of entangling pairings that were assigned to the 4-th layer. We reduced the count of entangling gates from 11 to 3, maintaining only [1,2], [5,6], and [9,10] in the QMI range of $[0.3, 0.4)$. The original set was composed by [0, 5], [1, 2], [1, 4], [2, 7], [3, 6], [4, 9], [5, 6], [5, 8], [6, 11], [7, 10], and [9, 10].

We notice that the $\#CNOT$ used in $QIDA^{CX}$ and $QIDA^{SO4}$ differs only for the number of correlators involved in the 1-st layer. This is due to the fact that for the $QIDA^{CX}$ ansatz, all the layers except the first are composed of the V-shape configurations of the layer, which doubles the count of CNOTs, while $QIDA^{SO4}$ counts two CNOTs for each gate used, including the first layer.

6.4.3 Heuristic ansätze comparison

We carried out the comparison between Multi-QIDA configuration and general Heuristic ansatz. A ladder arranged ansätze is defined by a layer of parametrized rotations $R_y(\theta)$, $N-1$ CNOTs in a top-down ordered configuration, followed by another set of parametrized rotations. Both layers of rotations are applied to all the N qubits. The number of repetitions of this sequence of operators is called *depth*. In this case, with the term depth, we refer to the number of entanglement block repetitions. In particular, we compared the Multi-QIDA approach with parametrized circuits having ladder fashion ansatz. We compared the number of CNOTs that are required to composed the two types of ansätze. In a situation in which the number of CNOTs of a multi-QIDA configuration can not be compared directly with the exact depth of the ladder circuit, we decided to confront ladders

that are composed of the previous and next depths. Thus, if the ladder contains fewer CNOTs than our approach, and the ansatz with depth increased by one, contains a higher number of CNOTs, then we take both ladders as a comparison. The number of CNOTs for ladder fashion circuits is defined as $(N-1)*d$, where N is the number of qubits, i.e. sites number, and d is the number of repetition of each layer with ladder displacement of the entangling gate i.e. the depth. We refer to the ladder ansätze as $(L)_d^{CX}$.

For the 3×4 Isotropic Hamiltonian model, the number of sites is 12, which translates to 11 entangling gates per layer. The ladder ansätze we used as comparison contain 44 CNOTs for $(L)_4^{CX}$, 55 CNOTs for $(L)_5^{CX}$, and 66 CNOTs for $(L)_6^{CX}$.

6.4.4 Performance Analysis

Here, we present the comparison between our Multi-QIDA approach and the heuristic ladder ansätze for one specific system that we have studied, the 3×4 Isotropic Heisenberg Model Hamiltonian. As already introduced in the previous section, we carried out the comparison between ladder fashion ansätze and Multi-QIDA method circuits, in both CNOTs and **SO**(4) parametrized gates.

- *Heuristic ladder-fashion ansätze*: For this type of ansätze, the results we obtained are very close between PQC at different depths. In particular, the two deeper circuits, with depths 5 and 6, present small variations in terms of mean result and best-performing run. The deeper circuit is the one that performs slightly better. For the absolute quantum error (AQE), the average is set to 84.37%, and the best-performing experiment reaches 90.36%, with a mean error from the best result of 5.99%. The relative quantum error (RQE) values are 57.17%, 73.58%, and 16.42%, corresponding to mean, best VQE, and deviation, in order. The ladder fashion circuits present on average a dispersion from the best-performing runs that is at least of 5.52% for the absolute quantum energy and at least 15.10% for the relative quantum energy. Looking at mean energy deviation error (MEAD), we can see that the best-performing ladder incurs an error of 0.4.

Lattice	*Ansatz*	E_{avg}	E_{best}	AQE_{avg}	RQE_{avg}	AQE_{best}	RQE_{best}	MED	$MAED$	$MRED$
3×4	$(L)_4^{CX}$	$-5.56(\pm 24)$	-5.924784	83.03	53.49	88.54	68.59	0.368738	5.51	15.10
3×4	$(L)_5^{CX}$	$-5.65(\pm 14)$	-6.067645	84.36	57.14	90.67	74.44	0.422450	6.31	17.30
3×4	$(L)_6^{CX}$	$-5.65(\pm 25)$	-6.046642	84.37	57.17	90.36	73.58	0.401	5.99	16.42
3×4	$(QIDA)^{CX}$	$-6.200065(\pm 1)$	-6.200066	92.65	79.87	92.65	79.87	$\sim$ 1e-6	$\sim$ 2e-5	$\sim$ 5e-5
3×4	$(QIDA)^{SO4}$	$-6.3612(\pm 42)$	-6.362179	95.061	86.46	95.08	86.51	0.000998	0.014918	0.040885

Table 6.2: Experimental results for the 3×4 Isotropic Heisenberg model Hamiltonian divided by ansatz layout. Results were obtained with $\#VQE = 50$ each, statevector simulation using Qiskit library, optimization algorithm BFGS with convergence tolerance 1e-6, and finesse ratio $\mu = 0.1$ for the layer selection. AQE_{avg} and RQE_{avg} are defined as the average performance metrics AQE_i and RQE_i over $\#VQE$ experiments. For the E_{avg}, the value in the braket is the standard deviation.

- *Multi-QIDA with CNOTS*: We increased the mean AQE to 92.65% and the best-performing run too, with a value of 92.65. The mean absolute energy deviation that we commit using this ansatz is 2e-5%. Consequently, the RQE presents the same behavior, 79.87% for the average RQE and best VQE, and an error of 5e-5%. For the $QIDA^{CX}$ ansatz, the RQE_{avg}

is increased by +22.70%, the RQE_{best} by +6.29%, with increased precision of seven orders of magnitude. In terms of MAED, we obtain an energy deviation of $1e$-6.

- *Multi-QIDA with SO(4)*: We obtained 95.061%, 95.08% and 0.149%, for average AQE, AQE_{best} and deviation, in order. The RQE results are instead 86.46%, 86.51, and 0.04% for mean RQE, best-experiment RQE and minimum relative deviation. In comparison to the ladders, for the $QIDA^{SO4}$, the enhancements that we obtained are +29.29% and +12.93%, for RQE_{avg} and RQE_{best}, respectively. The deviation from the minimum, $MRED$, is three orders of magnitude lower than the one obtained with the deeper heuristic ladder ansatz. We obtained a MEAD of $1e$-3 for this ansätze configuration.

In Table 6.2, we have collected all the results for this system configuration.

6.4.5 Convergence and precision

For a pictorial representation of the results shown in Table 6.2, a series of resuming violin plots can be found in Fig.6.7.

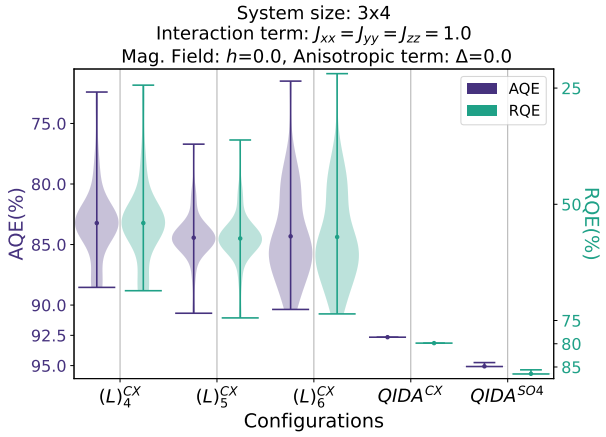


Figure 6.7: 3×4 VQE results. Comparison between AQE and RQE for the different ansätze settings.

Each violin represents the AQE and the RQE for a different ansätze configuration for the same system, tested in both ladder and QIDA ansatz settings. The lower tick represents the best VQE, and the higher tick represents the worst VQE. The black dots are associated to the average value of both AQE and RQE. The width of the violins represents the frequency of the experimental outcomes and they can be thought of as a rotated and smoothed histogram. From these plots, it is possible to see the difference in the performance between general heuristic ansätze and Multi-QIDA variational forms.

The ladder-fashion ansätze present average results that are generally distant from the minimum results, meaning that it is required to re-run the circuit different times before obtaining a satisfying result. Ladders can also incur into optimization paths that may converge to local minima that are far away from the best result that can be obtained with such a parametrized circuit. As you can see for the depth 4 ladder, only a few runs can reach similar results to the best-performing simulation for that configuration.

Multi-QIDA approaches instead guide the variational form in the right spot, providing good results with very high probability, thus, few shots are required to obtain these results. From Table

6.2, it is possible to see this behaviour noting that the difference between the average case and best-performing simulation is almost overlappable. In terms of energetic results and recovery of the highest correlation, $QIDA^{SO4}$ configuration provides the best results. Looking at the errors committed by the two Multi-QIDA configurations, the $QIDA^{CX}$ ansätze maintain the optimal convergence and lower MRED. The $QIDA^{SO4}$ ansätze present a slightly larger inaccuracy with an error that is three orders of magnitude bigger than the CNOT implementation.

6.4.6 Layer-wise VQE optimization routine

In this section, we describe the cost of the Multi-QIDA approach compared to the ladder ansätze in terms of evaluations of the cost function and dispersion of the optimization trajectory with respect to the best-performing and average VQE run. In Figure 6.8, we show the convergence plot for the 3×4 isotropic Heisenberg system. Here, we have collected all the optimization trajectories for each of the experiments. Each lighter-colored line represents one VQE simulation, while we plotted in a darker color, the best-performing VQE and average VQE values. For these plots, it is possible to see how the VQEs simulations with Multi-QIDA ansätze tend to be more compact in terms of dispersion from the best results. Thus, we can say that with high precision Multi-QIDA-ansätzes drive almost all the VQE simulations to follow the same path as the best-performing VQE run.

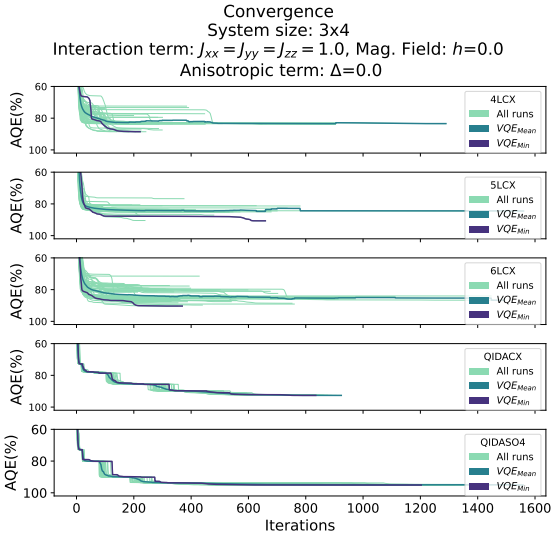


Figure 6.8: Representation of the optimization trajectories plotted by each VQE in terms of AQE. From top to bottom: The first three plots represent ladder-fashion ansätze, then the $QIDA^{CX}$ circuit, and finally, the **SO**(4) implementation.

In terms of the average number of iterations required for the different types of ansatz, it is clear that it is higher for the Multi-QIDA approach. This is due to the iterative optimization nature of the Multi-QIDA approach, against the standard VQE for the standard heuristic ladder ansatz which requires a single VQE routine. For the 3×4 isotropic Heisenberg Hamiltonian system, we obtained that the number of optimization VQE steps for the three ladders ansatz is 338, 427, and 583, respectively for depth 3, 4, and 5. For our Multi-QIDA ansatz, we reached 825 evaluations for the CNOT version and 1073 for the **SO**(4) implementation. The mean number of iterations required by the Multi-QIDA ansatze is at least double the evaluations required by the heuristic ladder ansätze.

From Figure 6.8, it is also possible to note that, the greater part of the optimization contribution, in terms of the number of iterations, is found in the last part of optimization, and during each layer re-optimization. This is because the final optimization, which is the re-optimization of the whole circuit, is the procedure that tries to optimize an already converged solution, and it is in general the longest contribution in the total number of iterations. In this situation, we could have implemented a different halting criterion for the optimization, which could have helped in reducing the time in which the optimization is almost converged. Using a fixed number of evaluations for the re-optimization routine or an additional threshold on the ΔE could help speed-up the whole procedure.

6.4.7 Summary

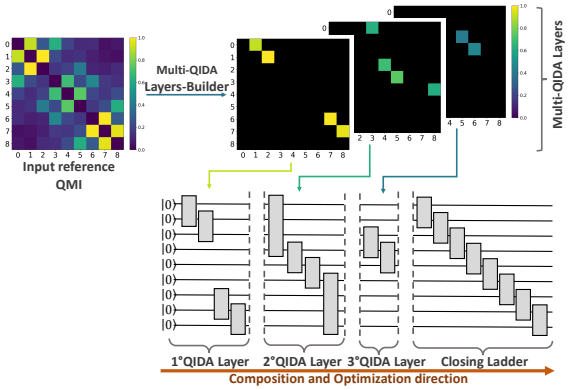
In this work, we proposed a state-preparation approach based on quantum mutual information extended from a previous work in which this approach is tested for molecular systems. Our approach aims to recover the correlation that is missed by the standalone QIDA method, introducing a multi-layer based ansatze. Our approach is surely more costly in terms of the number of evaluations of the cost function in which the convergence is reached, but, the error made on average is at least two order of magnitude lower than a standard heuristic ladder approach. Also, the correlation energy that is capture by our approach, for both $QIDA^{CX}$ and $QIDA^{SO4}$, is in general higher.

It is also worth noting that, differently from molecular systems, the reference state,i.e. the neél state, it could be very far away in energy from the groundstate energy. Generally, for the molecules, the Hartree-Fock reference state reaches almost 98% of the energy total energy, and post-Hartree-Fock algorithms as well as quantum algorithms, work in the remaining 2% of energy. For the systems studied in this work, the room of work is larger, with configurations in which the reference state is at 60% of the total energy. Thus, the improvement range is wider for spin systems in terms of RQE, which is a similar quantity to the *correlation energy*. In this term, further experiments on molecular systems are required to understand if this approach could be useful in quantum chemistry and how well it can perform on that type of system.

6.5 Conclusion

In this work, we introduced the Multi-Threshold Quantum Information Driven Ansatz (Multi-QIDA), a novel approach to enhance quantum simulations of strongly interactive lattice spin models. By extending the single-threshold QIDA method, Multi-QIDA leverages multiple thresholds of Quantum Mutual Information (QMI) to construct more efficient and compact quantum circuits. Our approach utilizes a layered structure, where each layer is built based on the QMI values of qubit pairs, enabling the inclusion of mid-high QMI value pairs and thereby capturing more correlations within the system.

The effectiveness of Multi-QIDA was validated through extensive benchmarking on various configurations of the Heisenberg model Hamiltonian. Our results demonstrated that Multi-QIDA outperforms traditional heuristic ansatz methods, achieving higher precision in ground state energy calculations while reducing computational complexity. Specifically, Multi-QIDA circuits showed a significant improvement in accuracy and convergence, with fewer iterations needed to reach optimal solutions compared to ladder-fashion heuristic ansatzes.



The iterative optimization routine of Multi-QIDA, which combines shallow and adaptable quantum circuits, proved to be an effective strategy in navigating the parameter space and mitigating issues such as barren plateaus. This iterative approach allowed for a more targeted exploration of the quantum landscape, leading to more reliable and consistent results.

Overall, the Multi-QIDA method presents a promising advancement in the field of quantum simulations for lattice spin models. Its ability to efficiently utilize quantum hardware and deliver high-precision results positions it as a valuable tool for future research in quantum computing and quantum chemistry. Further exploration and adaptation of this method to other quantum systems could potentially unlock new possibilities and applications in the realm of quantum simulations.

Conclusion

This PhD thesis focuses on building a bridge between classically computed and quantum computed quantum chemistry. Historically, the connection between these two fields has primarily been established through the DMRG algorithm, which is limited to systems with only a few hundred orbitals when applied to quantum chemistry. Similarly, the application of quantum information-theoretical quantities to enhance computational methods has been largely unexplored, with only a few exceptions in Tensor Network based approaches. This limited utilization also extends to theoretical analysis more broadly, where quantum information concepts have seen restricted application.

Therefore, this PhD work aims to expand scientific knowledge across these domains, effectively laying the foundation for connecting quantum chemistry and quantum information theory. This goal is achieved by developing tools that generalize and extend the applications of quantum information theory in quantum chemistry, while maintaining focus on their potential applications in quantum computing.

In Chapter 3 and in [98], we introduced the SparQ method, designed to efficiently calculate quantum information-theoretical quantities for sparse wavefunctions in quantum chemistry. Traditional approaches like DMRG typically perform poorly for chemical systems with hundreds of qubits due to the inherent non-locality of the electronic Hamiltonian. SparQ addresses this limitation by generalizing the trace operator across multiple fermionic-to-qubit mappings, thereby enabling access to large-scale quantum systems. By leveraging the natural sparsity present in most quantum chemistry wavefunctions, SparQ efficiently maps fermionic wavefunctions into qubit space using transformations such as Jordan-Wigner, significantly expanding the scope of quantum information theoretical analysis. To demonstrate practical applications, we tested SparQ on mutual information analysis of water molecules across different fermionic-to-qubit mappings and conducted entropy analysis for various active spaces in the benzene molecule. These applications highlight SparQ's capability to handle large-scale wavefunctions, opening new avenues for understanding correlations and behaviors in complex molecular systems.

Our second development, presented in Chapter 4 and based on [97], builds upon the SparQ method to investigate the nature of orbital-wise correlations in quantum chemistry—specifically distinguishing between quantum and classical correlations. Through computational techniques and theoretical frameworks developed for analyzing the results, we conducted numerical simulations on various medium-sized chemical systems using high-level post-Hartree-Fock methods. Our findings demonstrate that Natural Orbitals play a pivotal role in minimizing purely quantum von Neumann-only qubit-qubit correlations while retaining predominantly classical contributions. This result provides significant computational simplification for processing quantum information quantities in quantum chemistry and substantiates the widespread adoption of Natural Orbitals. Additionally,

it raises important questions regarding the true source of computational complexity in multi-body quantum chemistry problems and how Natural Orbitals might offer a pathway to reduce this complexity. This work establishes a foundation for further exploration of classical and quantum correlation components in quantum systems, potentially informing future computational strategies.

The third work, detailed in Chapter 5 and derived from [96], introduces a novel approach to constructing hardware-efficient ansätze for molecular systems. Our method, called Quantum Information Driven Ansatz (QIDA), utilizes mutual information derived from wavefunctions calculated with efficient classical methods like MP2 to develop initial ansätze. These ansätze can achieve correlation energies comparable to or better than traditional heuristic ladder-type ansätze while significantly reducing the CNOT gate count. To validate the QIDA method's effectiveness, we conducted statevector simulations across various molecular systems. The results demonstrate that as the number of qubits increases, the QIDA method substantially outperforms traditional approaches, confirming its scalability and robustness. Furthermore, we propose advancing this method through a modular approach to ansatz construction, using QIDA as an effective starting point.

This concept is further developed in the fourth main contribution of this thesis, the Multi-Threshold Quantum Information-Driven Ansatz method presented in Chapter 6. Although we specifically applied Multi-QIDA to strongly interacting lattice spin models, it provides a general framework for quantum simulations across multiple Quantum Mutual Information thresholds, generating highly efficient quantum circuits for target states that could serve as state-preparation algorithms for fault-tolerant refinement. When benchmarked against the Heisenberg model Hamiltonian, this method demonstrated clear advantages over traditional heuristic ansatz approaches, achieving higher accuracy with faster convergence and fewer iterations. The iterative optimization routine of Multi-QIDA combines shallow and adaptive quantum circuits to enable efficient exploration of vast parameter spaces while mitigating barren plateau issues. This approach enhances quantum hardware efficiency and provides a reliable method for exploring quantum landscapes. Its ability to reduce computational complexity also establishes a pathway for applying Multi-QIDA to future quantum chemistry research.

In summary, these four contributions represent significant advancements in bridging quantum chemistry simulations and quantum computing research. SparQ has expanded our computational toolkit for exploring large-scale quantum systems, while our study of orbital-wise correlations has opened new avenues for investigating classical and quantum components of these interactions. The development of QIDA and Multi-QIDA methods has introduced more efficient and scalable approaches to quantum circuit construction, reducing the computational resources required for quantum simulations. These contributions will prove invaluable for enhancing current capabilities and facilitating future innovations in both quantum chemistry and quantum computing. As quantum hardware continues to evolve, the methods and insights presented in this thesis will play a crucial role in expanding the possibilities of quantum simulations, ultimately leading to deeper understandings of complex molecular systems through more effective computational strategies.

List of Publications

Published

- D. Materia, L. Ratini, C. Angeli and L. Guidoni, *Quantum Information Driven Ansatz (QIDA): shallow-depth empirical quantum circuits from Quantum Chemistry*, J. Phys. Chem., (https://doi.org/10.1021/acs.jpca.4c03756), 2024.
- D. Materia, L. Ratini, C. Angeli and L. Guidoni, *Quantum Information reveals that orbital-wise correlation is essentially classical in Natural Orbitals*, The Journal of Chemical Physics, (10.1063/5.0220306), 2025.
- F. Tarocco, D. Materia, L. Ratini and L. Guidoni, *Compact Multi-Threshold Quantum Information Driven Ansatz For Strongly Interactive Lattice Spin Models*, Journal of Physics A: Mathematical and Theoretical, (10.1088/1751-8121/adc4a1), 2025.

Under peer-review

- D. Materia, L. Ratini and L. Guidoni, *Quantum information theory on sparse wavefunctions and applications for Quantum Chemistry*, 10.48550/arXiv.2408.02631, 2024

Bibliography

[1] https://quantumtech.blog/2022/10/20/quantum-computing-modalities-a-qubit-primer-revisite

[2] https://edu.itp.phys.ethz.ch/hs12/qsit/script2.pdf.

[3] https://metaphor.ethz.ch/x/2020/hs/401-3461-00L/sc/HilbertianBases.pdf.

[4] S. Aaronson and D. Gottesman. Improved simulation of stabilizer circuits. *Physical Review A*, 70:052328, 11 2004.

[5] G. Aleksandrowicz, T. Alexander, P. Barkoutsos, L. Bello, Y. Ben-Haim, D. Bucher, F. J. Cabrera-Hernández, J. Carballo-Franquis, A. Chen, C.-F. Chen, J. M. Chow, A. D. Córcoles-Gonzales, A. J. Cross, A. Cross, J. Cruz-Benito, C. Culver, S. D. L. P. González, E. D. L. Torre, D. Ding, E. Dumitrescu, I. Duran, P. Eendebak, M. Everitt, I. F. Sertage, A. Frisch, A. Fuhrer, J. Gambetta, B. G. Gago, J. Gomez-Mosquera, D. Greenberg, I. Hamamura, V. Havlicek, J. Hellmers, Łukasz Herok, H. Horii, S. Hu, T. Imamichi, T. Itoko, A. Javadi-Abhari, N. Kanazawa, A. Karazeev, K. Krsulich, P. Liu, Y. Luh, Y. Maeng, M. Marques, F. J. Martín-Fernández, D. T. McClure, D. McKay, S. Meesala, A. Mezzacapo, N. Moll, D. M. Rodríguez, G. Nannicini, P. Nation, P. Ollitrault, L. J. O'Riordan, H. Paik, J. Pérez, A. Phan, M. Pistoia, V. Prutyanov, M. Reuter, J. Rice, A. R. Davila, R. H. P. Rudy, M. Ryu, N. Sathaye, C. Schnabel, E. Schoute, K. Setia, Y. Shi, A. Silva, Y. Siraichi, S. Sivarajah, J. A. Smolin, M. Soeken, H. Takahashi, I. Tavernelli, C. Taylor, P. Taylour, K. Trabing, M. Treinish, W. Turner, D. Vogt-Lee, C. Vuillot, J. A. Wildstrom, J. Wilson, E. Winston, C. Wood, S. Wood, S. Wörner, I. Y. Akhalwaya, and C. Zoufal. Qiskit: An open-source framework for quantum computing, 1 2019.

[6] D. Aliverti-Piuri, K. Chatterjee, L. Ding, K. Liao, J. Liebert, and C. Schilling. What can quantum information theory offer to quantum chemistry? *Faraday Discussions*, 2024.

[7] K. Andersson, P.-A. Malmqvist, and B. O. Roos. Second-order perturbation theory with a complete active space self-consistent field reference function. *The Journal of Chemical Physics*, 96(2):1218–1226, 1992.

[8] C. Angeli, R. Cimiraglia, S. Evangelisti, T. Leininger, and J.-P. Malrieu. Introduction of n-electron valence states for multireference perturbation theory. *The Journal of Chemical Physics*, 114(23):10252–10264, 2001.

[9] F. Aquilante, J. Autschbach, A. Baiardi, S. Battaglia, V. A. Borin, L. F. Chibotaru, I. Conti, L. D. Vico, M. Delcey, I. F. Galván, N. Ferré, L. Freitag, M. Garavelli, X. Gong, S. Knecht,

E. D. Larsson, R. Lindh, M. Lundberg, P. Åke Malmqvist, A. Nenov, J. Norell, M. Odelius, M. Olivucci, T. B. Pedersen, L. Pedraza-González, Q. M. Phung, K. Pierloot, M. Reiher, I. Schapiro, J. Segarra-Martí, F. Segatta, L. Seijo, S. Sen, D. C. Sergentu, C. J. Stein, L. Ungur, M. Vacher, A. Valentini, and V. Veryazov. Modern quantum chemistry with [open]molcas. *The Journal of Chemical Physics*, 152:214117, 6 2020.

[10] A. Aspuru-Guzik, A. D. Dutoi, P. J. Love, and M. Head-Gordon. Simulated quantum computation of molecular energies. *Science*, 309:1704–1707, 9 2005.

[11] S. Attila and O. N. S. *Modern Quantum Chemistry: Introduction to Advanced Electronic Structure Theory*. Dover Publications, 2012.

[12] G. Barcza, O. Legeza, K. H. Marti, and M. Reiher. Quantum-information analysis of electronic states of different molecular structures. *Physical Review A*, 83:012508, 1 2011.

[13] S. Barison, F. Vicentini, and G. Carleo. An efficient quantum algorithm for the time evolution of parameterized circuits. *Quantum*, 5:512, 2021.

[14] P. K. Barkoutsos, J. F. Gonthier, I. Sokolov, N. Moll, and G. S. et al. Quantum algorithms for electronic structure calculations: Particle-hole hamiltonian and optimized wave-function expansions. *Physical Review A*, 98:1–14, 2018.

[15] R. J. Bartlett, S. A. Kucharski, and J. Noga. Alternative coupled-cluster ansätze ii. the unitary coupled-cluster method. *Chemical Physics Letters*, 155:133–140, 2 1989.

[16] J. Bauschlicher, Charles W. The construction of modified virtual orbitals (MVO's) which are suited for configuration interaction calculations. *The Journal of Chemical Physics*, 72(2):880–885, 07 2008.

[17] M. C. Bañuls, R. Blatt, J. Catani, A. Celi, J. I. Cirac, M. Dalmonte, L. Fallani, K. Jansen, M. Lewenstein, S. Montangero, C. A. Muschik, B. Reznik, E. Rico, L. Tagliacozzo, K. V. Acoleyen, F. Verstraete, U.-J. Wiese, M. Wingate, J. Zakrzewski, and P. Zoller. Simulating lattice gauge theories within quantum technologies. *The European Physical Journal D*, 74:165, 8 2020.

[18] F. Benfenati, G. Mazzola, C. Capecci, P. K. Barkoutsos, P. J. Ollitrault, I. Tavernelli, and L. Guidoni. Improved accuracy on noisy devices by non-unitary variational quantum eigensolver for chemistry applications, 2021.

[19] F. Benfenati, G. Mazzola, C. Capecci, P. K. Barkoutsos, P. J. Ollitrault, I. Tavernelli, and L. Guidoni. Improved accuracy on noisy devices by nonunitary variational quantum eigensolver for chemistry applications. *Journal of Chemical Theory and Computation*, 17:3946–3954, 2021.

[20] M. Bensberg and M. Reiher. Corresponding active orbital spaces along chemical reaction paths. *The Journal of Physical Chemistry Letters*, 14(8):2112–2118, 2023. PMID: 36802629.

[21] M. Bensberg and M. Reiher. Corresponding active orbital spaces along chemical reaction paths. *The Journal of Physical Chemistry Letters*, 14:2112–2118, 3 2023.

[22] D. J. Bernstein. *Introduction to post-quantum cryptography*, pages 1–14. Springer Berlin Heidelberg, 2019.

[23] G. Berthier, A. Daoudi, and J. Flament. Natural orbitals in the cipsi algorithm. *Journal of Molecular Structure: THEOCHEM*, 166:81–86, 6 1988.

[24] J. Biamonte, P. Wittek, N. Pancotti, P. Rebentrost, N. Wiebe, and S. Lloyd. Quantum machine learning. *Nature*, 549(7671):195–202, 2017.

[25] J. Biamonte, P. Wittek, N. Pancotti, P. Rebentrost, N. Wiebe, and S. Lloyd. Quantum machine learning. *Nature*, 549:195–202, 9 2017.

[26] C. Bloch. Sur la théorie des perturbations des états liée. *Nuclear Physics*, 6:329–347, 1958.

[27] K. Boguslawski and P. Tecmer. Orbital entanglement in quantum chemistry. *International Journal of Quantum Chemistry*, 115:1289–1295, 10 2015.

[28] K. Boguslawski and P. Tecmer. Erratum: Orbital entanglement in quantum chemistry. *International Journal of Quantum Chemistry*, 117(24):e25455, 2017.

[29] K. Boguslawski, P. Tecmer, P. W. Ayers, P. Bultinck, S. D. Baerdemacker, and D. V. Neck. Efficient description of strongly correlated electrons with mean-field cost. *Physical Review B*, 89:201106, 5 2014.

[30] K. Boguslawski, P. Tecmer, G. Barcza, O. Legeza, and M. Reiher. Orbital entanglement in bond-formation processes. *Journal of Chemical Theory and Computation*, 9(7):2959–2973, 2013. PMID: 26583979.

[31] S. Bordoni and S. Giagu. Convolutional neural network based decoders for surface codes. *Quantum Information Processing*, 22:151, 3 2023.

[32] M. Born and R. Oppenheimer. Zur quantentheorie der molekeln. *Annalen der Physik*, 389, 1927.

[33] S. Bravyi, D. Browne, P. Calpin, E. Campbell, D. Gosset, and M. Howard. Simulation of quantum circuits by low-rank stabilizer decompositions. *Quantum*, 3:181, 9 2019.

[34] S. Bravyi, J. M. Gambetta, A. Mezzacapo, and K. Temme. Tapering off qubits to simulate fermionic hamiltonians. 1 2017.

[35] S. B. Bravyi and A. Y. Kitaev. Fermionic quantum computation. *Annals of Physics*, 298:210–226, 2002.

[36] R. H. Byrd, P. Lu, J. Nocedal, and C. Zhu. A limited memory algorithm for bound constrained optimization. *SIAM Journal on Scientific Computing*, 16:1190–1208, 1995.

[37] L. Bytautas, G. E. Scuseria, and K. Ruedenberg. Seniority number description of potential energy surfaces: Symmetric dissociation of water, n2, c2, and be2. *The Journal of Chemical Physics*, 143, 9 2015.

[38] C. Cade, L. Mineh, A. Montanaro, and S. Stanisic. Strategies for solving the fermi-hubbard model on near-term quantum computers. *Physical Review B*, 102:235122, 12 2020.

[39] Z. Cai, R. Babbush, S. C. Benjamin, S. Endo, W. J. Huggins, Y. Li, J. R. McClean, and T. E. O'Brien. Quantum error mitigation. *Reviews of Modern Physics*, 95:045005, 12 2023.

[40] B. Cooper and P. J. Knowles. Benchmark studies of variational, unitary and extended coupled cluster methods. *The Journal of Chemical Physics*, 133:234102, 12 2010.

[41] O. Crawford, B. van Straaten, D. Wang, T. Parks, E. Campbell, and S. Brierley. Efficient quantum measurement of pauli operators in the presence of finite sampling error. *Quantum*, 5:385, 1 2021.

[42] L. Dagum and R. Menon. Openmp: an industry standard api for shared-memory programming. *IEEE Computational Science and Engineering*, 5:46–55, 1998.

[43] E. R. Davidson. Properties and uses of natural orbitals. *Reviews of Modern Physics*, 44:451–464, 7 1972.

[44] E. R. Davidson. Selection of the Proper Canonical Roothaan-Hartree-Fock Orbitals for Particular Applications. I. Theory. *The Journal of Chemical Physics*, 57(5):1999–2005, 09 1972.

[45] L. Ding, S. Knecht, and C. Schilling. Quantum information-assisted complete active space optimization (QICAS). *The Journal of Physical Chemistry Letters*, 14:11022–11029, 12 2023.

[46] L. Ding, S. Knecht, Z. Zimborás, and C. Schilling. Quantum correlations in molecules: from quantum resourcing to chemical bonding. *Quantum Science and Technology*, 8(1):015015, dec 2022.

[47] L. Ding, S. Mardazad, S. Das, S. Szalay, U. Schollwöck, Z. Zimborás, and C. Schilling. Concept of orbital entanglement and correlation in quantum chemistry. *Journal of Chemical Theory and Computation*, 17:79–95, 2021.

[48] L. Ding and C. Schilling. Correlation paradox of the dissociation limit: A quantum information perspective. *Journal of Chemical Theory and Computation*, 16:4159–4175, 7 2020.

[49] T. H. Dunning. Gaussian basis sets for use in correlated molecular calculations. i. the atoms boron through neon and hydrogen. *The Journal of Chemical Physics*, 90:1007–1023, 1 1989.

[50] C. Edmiston and K. Ruedenberg. Localized atomic and molecular orbitals. *Rev. Mod. Phys.*, 35:457–464, Jul 1963.

[51] D. J. Egger, C. Capecci, B. Pokharel, P. K. Barkoutsos, L. E. Fischer, L. Guidoni, and I. Tavernelli. A study of the pulse-based variational quantum eigensolver on cross-resonance based hardware. *Physical Review Research*, pages 1–13, 2023.

[52] D. J. Egger, C. Gambella, J. Marecek, S. McFaddin, and M. M. et al. Quantum computing for finance: State-of-the-art and future prospects. *IEEE Transactions on Quantum Engineering*, 1:1–24, 2021.

[53] F. A. Evangelista. Alternative single-reference coupled cluster approaches for multireference problems: The simpler, the better. *The Journal of Chemical Physics*, 134:224102, 6 2011.

[54] D. A. Fedorov, B. Peng, N. Govind, and Y. Alexeev. Vqe method: a short survey and recent developments. *Materials Theory*, 6:2, 1 2022.

[55] R. P. Feynman. Simulating physics with computers. *International Journal of Theoretical Physics*, 21(6):467–488, 1982.

[56] J. P. Finley and K. Hirao. Second-order many-body perturbation theory with non-canonical and non-orthogonal orbitals. *Chemical Physics Letters*, 328(1):51–59, 2000.

[57] L. E. Fischer, T. Dao, I. Tavernelli, and F. Tacchino. Dual-frame optimization for informationally complete quantum measurements. *Phys. Rev. A*, 109:062415, Jun 2024.

[58] M. Fishman, S. R. White, and E. M. Stoudenmire. Codebase release 0.3 for ITensor. *SciPost Phys. Codebases*, pages 4–r0.3, 2022.

[59] M. Fishman, S. R. White, and E. M. Stoudenmire. The ITensor Software Library for Tensor Network Calculations. *SciPost Phys. Codebases*, page 4, 2022.

[60] R. Fletcher. *Practical Methods of Optimization*. John Wiley & Sons, New York, NY, USA, second edition, 1987.

[61] J. M. Foster and S. F. Boys. Canonical configurational interaction procedure. *Rev. Mod. Phys.*, 32:300–302, Apr 1960.

[62] M. Ganzhorn, D. J. Egger, P. Barkoutsos, P. Ollitrault, G. Salis, N. Moll, M. Roth, A. Fuhrer, P. Mueller, S. Woerner, I. Tavernelli, and S. Filipp. Gate-efficient simulation of molecular eigenstates on a quantum computer. *Physical Review Applied*, 11, 2019.

[63] T. Giurgica-Tiron, Y. Hindy, R. Larose, A. Mari, and W. J. Zeng. Digital zero noise extrapolation for quantum error mitigation. pages 306–316. Institute of Electrical and Electronics Engineers Inc., 10 2020.

[64] H. R. Grimsley, S. E. Economou, E. Barnes, and N. J. Mayhall. An adaptive variational algorithm for exact molecular simulations on a quantum computer. *Nature Communications*, 10, 2019.

[65] B. Harrison, D. Nelson, D. Adamiak, and J. Whitfield. Reducing the qubit requirement of Jordan-Wigner encodings of N-mode, K-fermion systems from N to $\lceil \log_2 \binom{N}{K} \rceil$. 11 2022.

[66] D. R. Hartree. The wave mechanics of an atom with a non-coulomb central field. part ii. some results and discussion. *Mathematical Proceedings of the Cambridge Philosophical Society*, 24:111–132, 1928.

[67] W. W. Ho and T. H. Hsieh. Efficient variational simulation of non-trivial quantum states. *SciPost Phys.*, 6:029, 2019.

[68] M. R. Hoffmann and J. Simons. A unitary multiconfigurational coupled-cluster method: Theory and applications. *The Journal of Chemical Physics*, 88:993–1002, 1 1988.

[69] M. R. Hoffmann and J. Simons. A unitary multiconfigurational coupled-cluster method: Theory and applications. *The Journal of Chemical Physics*, 88:993, 8 1998.

[70] B. Huron, J. P. Malrieu, and P. Rancurel. Iterative perturbation calculations of ground and excited state energies from multiconfigurational zeroth-order wavefunctions. *The Journal of Chemical Physics*, 58:5745–5759, 6 1973.

[71] H. Hussain, M. B. Javaid, F. S. Khan, A. Dalal, and A. Khalique. Optimal control of traffic signals using quantum annealing. *Quantum Information Processing*, 19:312, 9 2020.

[72] F. Illas, J. Rubio, and J. M. Ricart. Approximate natural orbitals and the convergence of a second order multireference many-body perturbation theory (cipsi) algorithm. *The Journal of Chemical Physics*, 89:6376–6384, 11 1988.

[73] J. A. Jafri and J. L. Whitten. Iterative natural orbitals for configuration interaction using perturbation theory. *Theoretica Chimica Acta*, 44:305–313, 1977.

[74] Z. Jiang, A. Kalev, W. Mruczkiewicz, and H. Neven. Optimal fermion-to-qubit mapping via ternary trees with applications to reduced quantum states learning. *Quantum*, 4:276, 6 2020.

[75] P. Jordan and E. Wigner. Über das paulische Äquivalenzverbot. *Zeitschrift fur Physik*, 47:631–651, 9 1928.

[76] A. Kandala, A. Mezzacapo, K. Temme, M. Takita, M. Brink, J. M. Chow, and J. M. Gambetta. Hardware-efficient variational quantum eigensolver for small molecules and quantum magnets. *Nature*, 549:242–246, 2017.

[77] S. Keller, M. Dolfi, M. Troyer, and M. Reiher. An efficient matrix product operator representation of the quantum chemical hamiltonian. *The Journal of Chemical Physics*, 143, 12 2015.

[78] S. F. Keller and M. Reiher. Determining factors for the accuracy of dmrg in chemistry. *CHIMIA*, 68(4):200, Apr. 2014.

[79] R. A. Kendall, T. H. Dunning, and R. J. Harrison. Electron affinities of the first-row atoms revisited. systematic basis sets and wave functions. *The Journal of Chemical Physics*, 96:6796–6806, 5 1992.

[80] M. E. Khatib, G. L. Bendazzoli, S. Evangelisti, W. Helal, T. Leininger, L. Tenti, and C. Angeli. Beryllium dimer: A bond based on non-dynamical correlation. *The Journal of Physical Chemistry A*, 118:6664–6673, 8 2014.

[81] A. Khedkar and M. Roemelt. Active space selection based on natural orbital occupation numbers from n -electron valence perturbation theory. *Journal of Chemical Theory and Computation*, 15:3522–3536, 6 2019.

[82] A. A. Kirillov and A. D. Gvishiani. *Theorems and problems in functional analysis*. Springer Science & Business Media, 2012.

[83] A. Y. Kitaev. Quantum measurements and the abelian stabilizer problem, 1995.

[84] W. Kutzelnigg. Error analysis and improvements of coupled-cluster theory. *Theoretica Chimica Acta*, 80:349–386, 1991.

[85] C. Lanczos. An iteration method for the solution of the eigenvalue problem of linear differential and integral operators. *J. Res. Natl. Bur. Stand. B*, 45:255–282, 1950.

[86] M. Larocca, S. Thanasilp, S. Wang, K. Sharma, J. Biamonte, P. J. Coles, L. Cincio, J. R. McClean, Z. Holmes, and M. Cerezo. A review of barren plateaus in variational quantum computing, 2024.

[87] J. Lee, W. J. Huggins, M. Head-Gordon, and K. B. Whaley. Generalized unitary coupled cluster wave functions for quantum computation. *Journal of Chemical Theory and Computation*, 15:311–324, 2019.

[88] S. Lee, J. Lee, H. Zhai, Y. Tong, A. M. Dalzell, A. Kumar, P. Helms, J. Gray, Z.-H. Cui, W. Liu, M. Kastoryano, R. Babbush, J. Preskill, D. R. Reichman, E. T. Campbell, E. F. Valeev, L. Lin, and G. K.-L. Chan. Evaluating the evidence for exponential quantum advantage in ground-state quantum chemistry. *Nature Communications*, 14:1952, 4 2023.

[89] Legeza and J. Sólyom. Optimizing the density-matrix renormalization group method using quantum information entropy. *Physical Review B - Condensed Matter and Materials Physics*, 68:195116, 11 2003.

[90] D. S. Levine, D. Hait, N. M. Tubman, S. Lehtola, K. B. Whaley, and M. Head-Gordon. Casscf with extremely large active spaces using the adaptive sampling configuration interaction method. *Journal of Chemical Theory and Computation*, 16:2340–2354, 4 2020.

[91] X. Liu, A. Angone, R. Shaydulin, I. Safro, Y. Alexeev, and L. Cincio. Layer vqe: A variational approach for combinatorial optimization on noisy quantum computers. *IEEE Transactions on Quantum Engineering*, 3:1–20, 2022.

[92] S. Lloyd. Universal quantum simulators. *Science*, 273(5278):1073–1078, 1996.

[93] P. O. Löwdin and H. Shull. Natural orbitals in the quantum theory of two-electron systems. *Physical Review*, 101:1730, 1956.

[94] I. Magoulas and F. A. Evangelista. Unitary coupled cluster: Seizing the quantum moment. *The Journal of Physical Chemistry A*, 127:6567–6576, 8 2023.

[95] J.-P. Malrieu, P. Durand, and J.-P. Daudey. Intermediate hamiltonians as a new class of effective hamiltonians. *Journal of Physics A: Mathematical and General*, 18(5):809, 1985.

[96] D. Materia, L. Ratini, C. Angeli, and L. Guidoni. Quantum information driven ansatz (qida): Shallow-depth empirical quantum circuits from quantum chemistry. *The Journal of Physical Chemistry A*, 9 2024.

[97] D. Materia, L. Ratini, C. Angeli, and L. Guidoni. Quantum information reveals that orbital-wise correlation is essentially classical in natural orbitals, 2024.

[98] D. Materia, L. Ratini, and L. Guidoni. Quantum information theory on sparse wavefunctions and applications for quantum chemistry, 2024.

[99] J. R. McClean, J. Romero, R. Babbush, and A. Aspuru-Guzik. The theory of variational hybrid quantum-classical algorithms. *New Journal of Physics*, 18:023023, 2 2016.

[100] A. Menczer, M. van Damme, A. Rask, L. Huntington, J. Hammond, S. S. Xantheas, M. Ganahl, and Ö. Legeza. Parallel implementation of the density matrix renormalization group method achieving a quarter petaflops performance on a single dgx-h100 gpu node. *Journal of Chemical Theory and Computation*, 20(19):8397–8404, 2024. PMID: 39297788.

[101] M. Metcalf, N. P. Bauman, K. Kowalski, and W. A. D. Jong. Resource-efficient chemistry on quantum computers with the variational quantum eigensolver and the double unitary coupled-cluster approach. *Journal of Chemical Theory and Computation*, 16:6165–6175, 2020.

[102] A. Miller, A. Glos, and Z. Zimborás. Treespilation: Architecture- and state-optimised fermion-to-qubit mappings, 2024.

[103] A. Miller, Z. Zimborás, S. Knecht, S. Maniscalco, and G. García-Pérez. Bonsai algorithm: Grow your own fermion-to-qubit mappings. *PRX Quantum*, 4:030314, 8 2023.

[104] J. W. Mullinax and N. M. Tubman. Large-scale sparse wave function circuit simulator for applications with the variational quantum eigensolver. *The Journal of Chemical Physics*, 162(7):074114, 02 2025.

[105] C. Møller and M. S. Plesset. Note on an approximation treatment for many-electron systems. *Physical Review*, 46, 1934.

[106] C. Møller and M. S. Plesset. Note on an approximation treatment for many-electron systems. *Physical Review*, 46, 1934.

[107] F. Neese. Software update: the orca program system, version 4.0. *WIREs Computational Molecular Science*, 8, 1 2018.

[108] F. Neese, T. Petrenko, D. Ganyushin, and G. Olbrich. Advanced aspects of ab initio theoretical optical spectroscopy of transition metal complexes: Multiplets, spin-orbit coupling and resonance raman intensities. *Coordination Chemistry Reviews*, 251:288–327, 02 2007.

[109] M. A. Nielsen and I. L. Chuang. *Quantum Computation and Quantum Information*. Cambridge University Press, 2010.

[110] J. Olsen. The casscf method: A perspective and commentary. *International Journal of Quantum Chemistry*, 111:3267–3272, 11 2011.

[111] T. Parella-Dilmé, K. Kottmann, L. Zambrano, L. Mortimer, J. S. Kottmann, and A. Acín. Reducing entanglement with physically-inspired fermion-to-qubit mappings, 2023.

[112] D. Peral-García, J. Cruz-Benito, and F. J. García-Peñalvo. Systematic literature review: Quantum machine learning and its applications. *Computer Science Review*, 51:100619, 2024.

[113] A. Peruzzo, J. McClean, P. Shadbolt, M. H. Yung, X. Q. Zhou, P. J. Love, A. Aspuru-Guzik, and J. L. O'Brien. A variational eigenvalue solver on a photonic quantum processor. *Nature Communications*, 5, 7 2014.

[114] K. A. Peterson, D. E. Woon, and T. H. Dunning. Benchmark calculations with correlated molecular wave functions. iv. the classical barrier height of the h+h2→h2+h reaction. *The Journal of Chemical Physics*, 100:7410–7415, 5 1994.

[115] J. Pipek and P. G. Mezey. A fast intrinsic localization procedure applicable for and semiempirical linear combination of atomic orbital wave functions. *The Journal of Chemical Physics*, 90:4916–4926, 5 1989.

[116] M. Pitonak, P. Neogrady, J. Rezac, P. Jurecka, M. Urban, and P. Hobza. Benzene dimer: High-level wave function and density functional theory calculations. *Journal of Chemical Theory and Computation*, 4:1829–1834, 11 2008.

[117] M. J. D. Powell. *Advances in optimization and numerical analysis*, pages 51–67. Springer, 1994.

[118] B. P. Prascher, D. E. Woon, K. A. Peterson, T. H. Dunning, and A. K. Wilson. Gaussian basis sets for use in correlated molecular calculations. vii. valence, core-valence, and scalar relativistic basis sets for li, be, na, and mg. *Theoretical Chemistry Accounts*, 128:69–82, 1 2011.

[119] B. P. Pritchard, D. Altarawy, B. Didier, T. D. Gibson, and T. L. Windus. New basis set exchange: An open, up-to-date resource for the molecular sciences community. *Journal of Chemical Information and Modeling*, 59:4814–4820, 11 2019.

[120] L. Ratini, C. Capecci, F. Benfenati, and L. Guidoni. Wave function adapted hamiltonians for quantum computing. *Journal of Chemical Theory and Computation*, 18:899–909, 2022.

[121] L. Ratini, C. Capecci, and L. Guidoni. Optimization strategies in wahtor algorithm for quantum computing empirical ansatz: a comparative study. *Electronic Structure*, 5:045006, 12 2023.

[122] L. Ratini, C. Capecci, and L. Guidoni. Natural orbitals and sparsity of quantum mutual information. *Journal of Chemical Theory and Computation*, 20(9):3535–3542, 2024.

[123] A. G. Rattew, S. Hu, M. Pistoia, R. Chen, and S. Wood. A domain-agnostic, noise-resistant, hardware-efficient evolutionary variational quantum eigensolver. *Preprint*, 2020.

[124] A. G. Rattew, S. Hu, M. Pistoia, R. Chen, and S. Wood. A domain-agnostic, noise-resistant, hardware-efficient evolutionary variational quantum eigensolver, 2020.

[125] J. Rissler, R. M. Noack, and S. R. White. Measuring orbital interaction using quantum information theory. *Chemical Physics*, 323:519–531, 2006.

[126] J. Romero, R. Babbush, J. R. McClean, C. Hempel, P. J. Love, and A. Aspuru-Guzik. Strategies for quantum computing molecular energies using the unitary coupled cluster ansatz. *Quantum Science and Technology*, 4:014008, 2018.

[127] J. Romero, R. Babbush, J. R. McClean, C. Hempel, P. J. Love, and A. Aspuru-Guzik. Strategies for quantum computing molecular energies using the unitary coupled cluster ansatz. *Quantum Science and Technology*, 4:1–18, 2019.

[128] J. J. Sakurai and J. Napolitano. *Modern Quantum Mechanics*. Cambridge University Press, 9 2017.

[129] U. Schollwöck. The density-matrix renormalization group. *Rev. Mod. Phys.*, 77:259–315, 2005.

[130] U. Schollwöck. The density-matrix renormalization group in the age of matrix product states. *Annals of Physics*, 326(1):96–192, 2011. January 2011 Special Issue.

[131] C. E. Shannon. A mathematical theory of communication. *Bell System Technical Journal*, 27:379–423, 7 1948.

[132] P. Shor. Algorithms for quantum computation: discrete logarithms and factoring. In *Proc. 35th Annu. Symp. Found. Comput. Sci.*, pages 124–134. IEEE Comput. Soc. Press, 1994.

[133] P. W. Shor. Polynomial-time algorithms for prime factorization and discrete logarithms on a quantum computer. *SIAM Journal on Computing*, 26(5), 1997.

[134] K. Sivalingam, M. Krupicka, A. A. Auer, and F. Neese. Comparison of fully internally and strongly contracted multireference configuration interaction procedures. *The Journal of Chemical Physics*, 145(5):054104, 08 2016.

[135] I. O. Sokolov, P. K. Barkoutsos, P. J. Ollitrault, D. Greenberg, J. Rice, M. Pistoia, and I. Tavernelli. Quantum orbital-optimized unitary coupled cluster methods in the strongly correlated regime: Can quantum algorithms outperform their classical equivalents? *Journal of Chemical Physics*, 152:1–18, 2020.

[136] C. J. Stein and M. Reiher. Automated selection of active orbital spaces. *Journal of Chemical Theory and Computation*, 12:1760–1771, 2016.

[137] C. J. Stein and M. Reiher. autocas: A program for fully automated multiconfigurational calculations. *Journal of Computational Chemistry*, 40:2216–2226, 2019.

[138] Q. Sun, X. Zhang, S. Banerjee, P. Bao, M. Barbry, N. S. Blunt, N. A. Bogdanov, G. H. Booth, J. Chen, Z.-H. Cui, J. J. Eriksen, Y. Gao, S. Guo, J. Hermann, M. R. Hermes, K. Koh, P. Koval, S. Lehtola, Z. Li, J. Liu, N. Mardirossian, J. D. McClain, M. Motta, B. Mussard, H. Q. Pham, A. Pulkin, W. Purwanto, P. J. Robinson, E. Ronca, E. R. Sayfutyarova, M. Scheurer, H. F. Schurkus, J. E. T. Smith, C. Sun, S.-N. Sun, S. Upadhyay, L. K. Wagner, X. Wang, A. White, J. D. Whitfield, M. J. Williamson, S. Wouters, J. Yang, J. M. Yu, T. Zhu, T. C. Berkelbach, S. Sharma, A. Y. Sokolov, and G. K.-L. Chan. Recent developments in the PySCF program package. *The Journal of Chemical Physics*, 153(2):024109, 2020.

[139] T.-P. Sun, Z.-Y. Chen, Y.-J. Wang, C. Xue, H.-Y. Liu, X.-N. Zhuang, X.-F. Xu, Y.-C. Wu, and G.-P. Guo. Sparqsim: Simulating scalable quantum algorithms via sparse quantum state representations, 2025.

[140] P. R. Surján. *Second quantized approach to quantum chemistry: an elementary introduction.* Springer Science & Business Media, 2012.

[141] sWilliam J. Hunt and W. A. Goddard. Excited states of h2o using improved virtual orbitals. *Chemical Physics Letters*, 3(6):414–418, 1969.

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

[143] F. Tarocco, D. Materia, L. Ratini, and L. Guidoni. Compact multi-threshold quantum information driven ansatz for strongly interactive lattice spin models. *Journal of Physics A: Mathematical and Theoretical*, 2025.

[144] K. Temme, S. Bravyi, and J. M. Gambetta. Error mitigation for short-depth quantum circuits. *Physical Review Letters*, 119:1–15, 2017.

[145] N. V. Tkachenko, J. Sud, Y. Zhang, S. Tretiak, P. M. Anisimov, A. T. Arrasmith, P. J. Coles, L. Cincio, and P. A. Dub. Correlation-informed permutation of qubits for reducing ansatz depth in the variational quantum eigensolver. *PRX QUANTUM*, 2:20337, 2021.

[146] F. Vatan and C. Williams. Optimal quantum circuits for general two-qubit gates. *Physical Review A*, 69(3), 2004.

[147] V. von Burg, G. H. Low, T. Häner, D. S. Steiger, and M. R. et al. Quantum computing enhanced computational catalysis. *Physical Review Research*, 3, 2021.

[148] J. Watrous. *The Theory of Quantum Information.* Cambridge University Press, 4 2018.

[149] S. R. White. Density matrix formulation for quantum renormalization groups. *Phys. Rev. Lett.*, 69:2863–2866, 1992.

[150] S. R. White. Density-matrix algorithms for quantum renormalization groups. *Phys. Rev. B*, 48:10345–10356, 1993.

[151] J. D. Whitfield, J. Biamonte, and A. Aspuru-Guzik. Simulation of electronic structure hamiltonians using quantum computers. *Molecular Physics*, 109:735–750, 3 2011.

[152] R. Wiersema, C. Zhou, Y. de Sereville, J. F. Carrasquilla, Y. B. Kim, and H. Yuen. Exploring entanglement and optimization within the hamiltonian variational ansatz. *PRX Quantum*, 1:020319, Dec 2020.

[153] A. K. Wilson, T. van Mourik, and T. H. Dunning. Gaussian basis sets for use in correlated molecular calculations. vi. sextuple zeta correlation consistent basis sets for boron through neon. *Journal of Molecular Structure: THEOCHEM*, 388:339–349, 12 1996.

[154] S. Wouters and D. V. Neck. The density matrix renormalization group for ab initio quantum chemistry. *The European Physical Journal D*, 68:272, 9 2014.

[155] R. Xia and S. Kais. Qubit coupled cluster singles and doubles variational quantum eigensolver ansatz for electronic structure calculations. *Quantum Science and Technology*, 6:1–13, 2020.

[156] Y. S. Yordanov, V. Armaos, C. H. W. Barnes, and D. R. M. Arvidsson-Shukur. Qubit-excitation-based adaptive variational quantum eigensolver. *Communications Physics*, 4:228, 10 2021.

[157] T. Zelevinsky, A. F. Izmaylov, and A. N. Alexandrova. Physical chemistry of quantum information science. *The Journal of Physical Chemistry A*, 127(49):10357–10359, 2023. PMID: 38093691.

[158] B. Zeng, X. Chen, D.-L. Zhou, and X.-G. Wen. *Quantum Information Meets Quantum Matter*. Springer New York, 2019.

[159] Z.-J. Zhang, T. H. Kyaw, J. S. Kottmann, M. Degroote, and A. Aspuru-Guzik. Mutual information-assisted adaptive variational quantum eigensolver. *Quantum Science and technology*, 6(3):035001, 2021.

[160] J. Zheng, X. Xu, and D. G. Truhlar. Minimally augmented karlsruhe basis sets. *Theoretical Chemistry Accounts*, 128:295–305, 2 2011.

Acknowledgements

The main author acknowledge funding from the MoQS program, founded by the European Union's Horizon 2020 research and innovation under the Marie Skłodowska-Curie grant agreement number 955479.

The other collaborating authors acknowledge funding from Ministero dell'Istruzione dell'Università e della Ricerca (PON R & I 2014-2020), from National Centre for HPC, Big Data and Quantum Computing - PNRR Project, funded by the European Union - Next Generation EU and finally from the Ministero dell'Università e della Ricerca (MUR) under the Project PRIN 2022 number 2022W9W423 through the European Union Next Generation EU.