Quantum information theory on sparse wavefunctions and applications for Quantum Chemistry

Davide Materia[1,2], Leonardo Ratini[1,3], and Leonardo Guidoni [*1]

[1] Dipartimento di Scienze Fisiche e Chimiche, Università degli Studi dell'Aquila, Coppito, L'Aquila, Italy
[2]Dipartimento di Ingegneria e Scienze dell'Informazione e Matematica, Università degli Studi dell'Aquila, Coppito, L'Aquila, Italy
[3]Dipartimento di Scienze Matematiche, Fisiche e Informatiche, Università degli Studi di Parma, Parma, Italy

August 6, 2024

Abstract

In recent years Quantum Computing prominently entered in the field of Computational Chemistry, importing and transforming computational methods and ideas originally developed within other disciplines, such as Physics, Mathematics and Computer Science into algorithms able to estimate quantum properties of atoms and molecules on present and future quantum devices. An important role in this contamination process is attributed to Quantum Information techniques, having the twofold role of contributing to the analysis of electron correlation and entanglements and guiding the construction of wavefunction variational ansatzes for the Variational Quantum Eigensolver technique.

This paper introduces the tool SparQ (Sparse Quantum state analysis), designed to efficiently compute fundamental quantum information theory observables on post-Hartree-Fock wavefunctions sparse in their definition space. The core methodology involves mapping fermionic wavefunctions to qubit space using fermionic-to-qubits transformations and leveraging the sparse nature of these wavefunctions to evaluate observables and properties of the wavefunction.

The effectiveness of SparQ is validated by analyzing the mutual information matrices of wavefunctions for the water molecule and the total entropy of $\sim 10^2$ qubits describing the benzene moleculehighlights. This way its capability to handle large-scale quantum systems, limited mainly by the capabilities of quantum chemical methods used to retrieve the wavefunctions. The results indicate that quantum information theoretical analysis, so far limited to traditional tensor network methods and study of transition operators, can be applied to all post-Hartree-Fock wavefuncions, extending their applications to larger and more complex chemical systems.

1

1 Introduction

In recent years, the rapid growth of Quantum Computing (QC)[1] has highlighted the relevance of quantum information theory tools within computational sciences.[2;3;4;5] 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[6;7], this exchange has enabled for existing algorithms[8;9;10;11;12] to benefit from the inclusion of chemical concepts, exemplified by the UCCSD ansatz[13;14;15;16;17] for variational quantum chemistry algorithms. Additionally, advancements in quantum information theory have benefited classical quantum chemistry by enhancing the understanding of chemical properties[18;19;20] and introducing tools that focus computational efforts where they are most needed.[21]

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)[22;23;24;25;26] algorithm has proven to be a highly capable and versatile tool[27;28], 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[29;30]. Furthermore, given the inspiration from which this work arises[31], it falls directly in line with the mentioned concept of bringing QC and quantum computing closer.

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.

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 4 and forward. In quantum chemistry, the MPS-based DMRG[26] 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 $O(N^3\chi^2) + O(N^2\chi^3)$[32]. 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 of second order (MP2)[33]), which renders possible calculations on up to thousands of spin orbitals (~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,[34;35;19] 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 follows:

$$|\psi\rangle_{\mathcal{F}} = \sum_{\mathbf{i}=\{0,1\}^N} c_{\mathbf{i}} |\mathbf{i}\rangle_{\mathcal{F}} \qquad (1)$$

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[33], Configuration Interaction (CI)[7], Coupled Cluster (CC), and their multireference variants[36;37;38;39], employ relatively few states out of the exponentially many in (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 (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.

2.1 Encoding of fermionic wavefunctions

Beginning with the wavefunction definition in (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 (1), $\mathbf{i}$ will represent the binary string with ones indicating occupied fermionic modes. This means that, for a system of N fermionic modes,

$$a_{\mathbf{i}}^\dagger = \prod_{j=1}^N \left(a_j^\dagger\right)^{i_j} \qquad i_j = 0, 1 \qquad j = 1, \ldots, N \quad (2)$$

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

$$|\psi\rangle = \sum_{\mathbf{i}=\{0,1\}^N} c_{\mathbf{i}} a_{\mathbf{i}}^\dagger |\varnothing\rangle \qquad (3)$$

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

$$\{a_i, a_j\} = \{a_{q,i}, a_{q,j}\}$$
$$\{a_i^\dagger, a_j^\dagger\} = \{a_{q,i}^\dagger, a_{q,j}^\dagger\} \qquad (4)$$
$$\{a_i, a_j^\dagger\} = \{a_{q,i}, a_{q,j}^\dagger\} = \delta_{i,j}\mathbb{1}$$

In[30] 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

$$\mathcal{K}(|\varnothing\rangle) = |\mathbf{0}\rangle \tag{5}$$

$$\mathcal{K}(a_j^\dagger) = a_{q,j}^\dagger = \frac{1}{2}(\gamma_q^{2j-1} - i\gamma_q^{2j}) \tag{6}$$

$$\mathcal{K}(a_j) = a_{q,j} = \frac{1}{2}(\gamma_q^{2j-1} + i\gamma_q^{2j}) \tag{7}$$

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:

$$\gamma_q^l = \bigotimes_{m=1}^{N} \sigma_{l,m} \tag{8}$$

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

$$\mathcal{H}(a_{q,j}^\dagger) = \mathcal{H}(\gamma_q^{2j}, \gamma_q^{2j-1}) = \sum_{m=1}^{N} \delta_{\sigma_{2j,m}, \sigma_{2j-1,m}} \tag{9}$$

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 [30]. 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 A.

2.2 Jordan Wigner mapping

The Jordan-Wigner[29] mapping, 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:

$$\sigma_{l,m} = \begin{cases} I_m & \text{if } \lfloor\frac{l}{2}\rfloor > m \\ X_m & \text{if } \lfloor\frac{l+1}{2}\rfloor = m \\ Y_m & \text{if } \lfloor\frac{l}{2}\rfloor = m \\ Z_m & \lfloor\frac{l}{2}\rfloor - m \end{cases} \tag{10}$$

Using this definition, Eq. (7) assumes the following form:

$$a_{q,j}^\dagger = \bigotimes_{m=1}^{N} \epsilon_{j,m} \quad with$$

$$\epsilon_{j,m} = \begin{cases} \sigma_{2j,m}, & \text{if } \sigma_{2j,m} = \sigma_{2j+1,m} \\ \frac{1}{2}(\sigma_{2j,m} - i\sigma_{2j+1,m}), & \text{otherwise} \end{cases} \tag{11}$$

This allows us to rewrite a single SD of Eq. (3) as

$$\mathcal{K}(a_{\mathbf{i}}^\dagger) = a_{q,\mathbf{i}}^\dagger = \bigotimes_{m=1}^{N} \prod_{l=1}^{N} (\epsilon_{l,m})^{\mathbf{i}_l} \tag{12}$$

We can now rewrite directly Eq. (3) as

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

$$= \sum_{\mathbf{i}=\{0,1\}^N} c_{\mathbf{i}} \left(\bigotimes_{m=1}^{N} \prod_{l=1}^{N} (\epsilon_{l,m})^{\mathbf{i}_l} \right) |\mathbf{0}\rangle \tag{13}$$

Now, to pave the way for the usage of wavefunctions of this form in the following section 4, we write the wavefunction as follow

$$|\psi\rangle = \sum_{\mathbf{i}=\{0,1\}^N} c_{\mathbf{i}} \left(\bigotimes_{m=1}^{N} |\psi\rangle_{\mathbf{i},m} \right) = \sum_{\mathbf{i}=\{0,1\}^N} c_{\mathbf{i}} |\psi\rangle_{\mathbf{i}} \tag{14}$$

where

$$|\psi\rangle_{\mathbf{i},m} = \epsilon_{\mathbf{i},m} |0\rangle_m, \quad \epsilon_{\mathbf{i},m} = \prod_{l=1}^{N} (\epsilon_{l,m})^{\mathbf{i}_l} \tag{15}$$

4

$$and \quad |\psi\rangle_{\mathbf{i}} = \bigotimes_{m=1}^{N} |\psi\rangle_{\mathbf{i},m} \qquad (16)$$

our ket is now written as a sum of tensor products of single-qubit wavefunctions $|\psi\rangle_{i,m}$. The Eq. (14) 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 wave function.

One disclaimer, for the simple case of the Jordan Wigner 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. (1) to a binary expression over the qubit space, so that the following is true:

$$\sum_{\mathbf{i}} c_{\mathbf{i}} |\psi\rangle_{\mathbf{i}} = \sum_{\mathbf{i}} k_{\mathbf{i}} |\mathbf{i}\rangle \qquad (17)$$

where $|\mathbf{i}\rangle$ is binary ket. On the other hand, this is equivalent to realizing that in (11) and (15) no superposition is created by any $\epsilon_{l,m} |0\rangle_m$, as it can be seen by noting that:

$$\epsilon_{\mathbf{i},m} |0\rangle_m = \begin{cases} |0\rangle_m & \text{if } \mathbf{i}_m = 0 \\ \pm |1\rangle_m & \text{if } \mathbf{i}_m = 1 \end{cases} \qquad (18)$$

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 A).

3   Quantum Information

The primary objective of this work is to develop an efficient method for computing information-theoretical quantities for non-MPS wavefunctions. To achieve this, we must first evaluate the general complexity of such analyses.

In quantum information theory, analyses typically occur in a Hilbert space $\mathscr{H}$, where each vector represents a wavefunction $|\psi\rangle_{\mathscr{H}}$. In quantum computing, this Hilbert space consists of multiple separable two-dimensional spaces known as *qubits*. The entire space of N qubits $\mathscr{H}$ is

$$\mathscr{H} = \bigotimes_{l}^{N} \mathscr{H}_l \qquad (19)$$

with a total dimension of $\mathscr{H}$ is 2^{N} [40].

Given a wavefunction $|\phi\rangle \in \mathscr{H}$, its Density Matrix (DM) can defined over the space of *Endomorphisms* of $\mathscr{H}$, $End(\mathscr{H})$, such that

$$\rho_\phi = |\phi\rangle\langle\phi| \qquad (20)$$

This state description captures general quantum or statistical properties, but it requires a quadratic cost in terms of the components of the wavefunction ϕ. Given the exponential dimension of the space in relation to the number of qubits, analyzing the properties of wavefunctions and density matrices within this space is impractical.

To reach the goal, there is a need to make use of the sparse property of the wavefunctions of chemical systems discussed in section 2.

For a bipartite quantum system with Hilbert spaces $\mathscr{H}_A$ and $\mathscr{H}_B$, the partial trace over subsystem B of the density matrix ρ_{AB} is defined as:

$$\rho_A = \text{Tr}(\rho_{AB})_B = \sum_i {}_B\langle i|\rho_{AB}|i\rangle_B \qquad (21)$$

Where $|i\rangle_B$ are any orthonormal basis of the Hilbert space B. The primary use of the traced density matrix is to calculate the Von Neumann (or quantum) mutual information. [41]

Definition 1 *Consider two Hilbert spaces $\mathscr{H}_A$ and $\mathscr{H}_B$, and let $\rho_{AB} \in \mathscr{H}_A \otimes \mathscr{H}_B$ be a density matrix. The quantum mutual information I of ρ_{AB} is given by:*

$$I(A,B) = S(A) + S(B) - S(A,B) \qquad (22)$$

where S represents the von Neumann entropy defined

as

$$S(A, B) = -\operatorname{tr}(\rho_{AB} \log(\rho_{AB}))$$
$$S(A) = -\operatorname{tr}(\rho_A \log(\rho_A)) \qquad (23)$$
$$\rho_A = \operatorname{tr}_B(\rho_{AB})$$

As described in Eq. (23), for an N-partite Hilbert space $\mathscr{H}_N$ and a quantum state expressed by $\rho \in \mathscr{H}_N$, the elements $I(i, j)$ of the mutual information matrix I are obtained by tracing out all subsystems except i, j. In our use case, the subsystems consist of N spin-orbitals. Notably, the matrix is inherently symmetric, with zero values along the diagonal (by convention, as diagonal values are undefined).

The mutual information defined in Eq. (22) has been used in the past[42;21;20] for its flexibility in describing correlation and for its further ability to discern high levels of entanglement.

In the following we will also consider the classical counterpart, the Shannon entropy, which will be useful as a comparison to the von Neumann. Shannon entropy is calculated by evolving the density matrix through a measuring channel as for the following equation:

$$p(i) = \langle i| \rho |i\rangle$$
$$\mathcal{M}(\rho) = \sum_i p(i) |i\rangle \langle i| \qquad (24)$$

This channel collapses the wavefunction rendering it a classical statistical distribution with probabilities $p(i)$ over the space $|i\rangle$, which in our case is the computational basis.

4   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.

$$|\psi\rangle = \sum_j \lambda_j \left(\bigotimes_{m=1}^N |\psi\rangle_{j,m} \right) \qquad (25)$$

The first kind of wavefunction is the one expressed by Eq. (25), 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 (26), 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.

$$|\psi\rangle = \sum_{\mathbf{i}=\{0,1\}^N} k_{\mathbf{i}} |\mathbf{i}\rangle \qquad (26)$$

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. (25) 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 (25), 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 (26). 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. (25).

Beginning with the first of these two notations, we have shown in appendix A that any wavefunction can be encoded as Eq. (25). For fermionic wavefunctions

under the Jordan-Wigner mapping, this reduces to equation 14, with λ_j becoming the $c_{\mathbf{i}}$ coefficients as in (14).

4.1 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

$$\rho = |\psi\rangle\langle\psi| \tag{27}$$

that, by using Eq. 25, becomes

$$|\psi\rangle\langle\psi| =$$
$$= \sum_{j,j'} \lambda_j \lambda_{j'}^* \left(\bigotimes_{m=1}^{N} (|\psi\rangle_{j,m} \langle\psi|_{j',m}) \right) \tag{28}$$

4.2 Partial Trace

We now have the tools required to reduce the dimensionality of the qubit space starting from Eq. (28). The reduced density matrix corresponding to the qubit k can be obtained as follows:

$$\rho_k = tr(\rho)_{1,\cdots k-1,k+1,\cdots,N} = \sum_{j,j'} \lambda_j \lambda_{j'}^*$$
$$tr \left(\bigotimes_{m=1,m\neq k}^{N} (|\psi\rangle_{j,m} \langle\psi|_{j',m}) \right) \cdot |\psi\rangle_{j,k} \langle\psi|_{j',k} \tag{29}$$

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 $O(\chi^2 N)$

4.3 Observables

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

$$O = \sum_k o_k \bigotimes_{m=1}^{N} O_{k,m} \tag{30}$$

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

$$\langle\psi| O |\psi\rangle = tr(|\psi\rangle O \langle\psi|) =$$
$$= tr \left(\sum_{j,j',k} \lambda_j \lambda_{j'}^* o_k \left(\bigotimes_{m=1}^{N} (|\psi\rangle_{j,m} \langle\psi|_{j',m} O_{k,m}) \right) \right) \tag{31}$$

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. (29), however, there is now also a multiplicative cost in the number of operators defining O.

4.4 Direct tracing

The partial trace equation introduced in (29) requires evaluating all possible combinations of two states ($\sum_{j,j'}$) in the outer product defining the density matrix (27), significantly limiting the number of states χ of the wavefunction (25) 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. (26). 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. (26), 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. (26) 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:

$$(|\mathbf{i}\rangle)_A = |\mathbf{i}\rangle_A \qquad (32)$$

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 $|\mathbf{i}\rangle = |\mathbf{i}\rangle_A \otimes |\mathbf{i}\rangle_B$.

With this notation in mind, we can now retake Eq. (21) 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 $|\psi\rangle$ (26) 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 (28), 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 (27) 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.

$$\rho_A = tr(|\psi\rangle\langle\psi|)_{\overline{A}} \qquad (33)$$

To obtain a better scaling, and therefore exploiting the linearity of (21), 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. (29) 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:

1. We start by doing an iteration over all the binary strings in $|\psi\rangle$ (26)(line 3), during the iterations

---

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_ρ_A ← #Keys(γ_A);          // ← Dim(H_A)
11 allocate ρ_A(n_ρ_A, n_ρ_A);   // 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 (26)
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
```

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. (21), 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 $11-25$)

4.4.1 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 strategyy using the particle-hole formalism explained in Appendix C, 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.

This method can be generalized to partial traces of mixed density matrices which cannot be written as (33), 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.

5 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. (22), as it has long been an

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[43;44]	184	Direct	$1.2 \cdot 10^5$
C_6H_6	C —— H distance, 1.085 C —— C distance, 1.390	ccpvdz[43]	228	Direct	$2.5 \cdot 10^5$

Table 1: Summary of tested molecular systems together with the computational details. The parameter χ is the number of SD.

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 (21) has been written in $C++$ to guarantee easy portability to parallelizing platforms such as OpenMP[45] while the sparse wavefunctions are all relative to CISD calculations found with the Pyscf[46] 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[43;44] 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[43] basis-set. All the data listed here is also summarized by Tab. 1.

6 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.

a)

b)
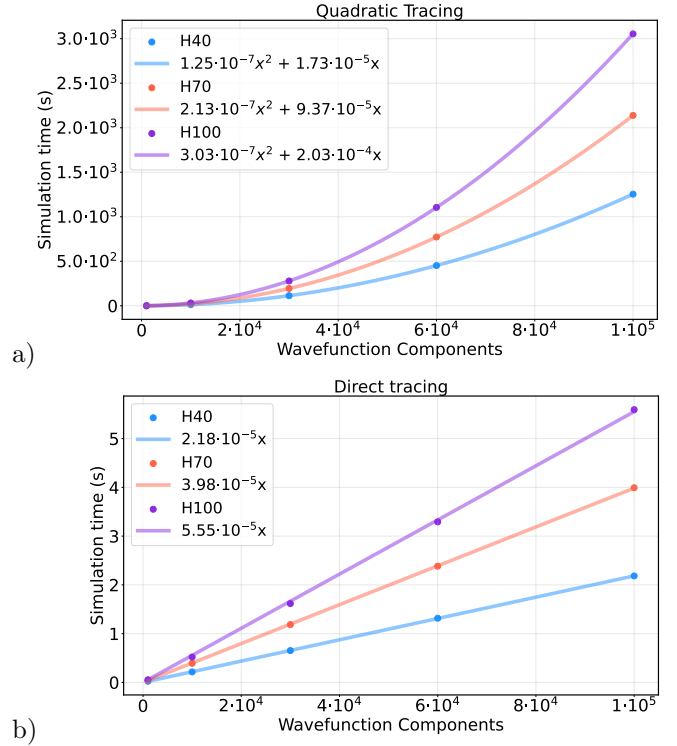
Figure 1: Time requirements for the partial trace from N qubits to one qubits over different values of number of components of the wavefunction a) Quadratic tracing method of Eq. (29), b) Tracing method of pseudocode 1 with linear scaling. As shown in 1 each chain H_x is expressed with $4y$ qubits.
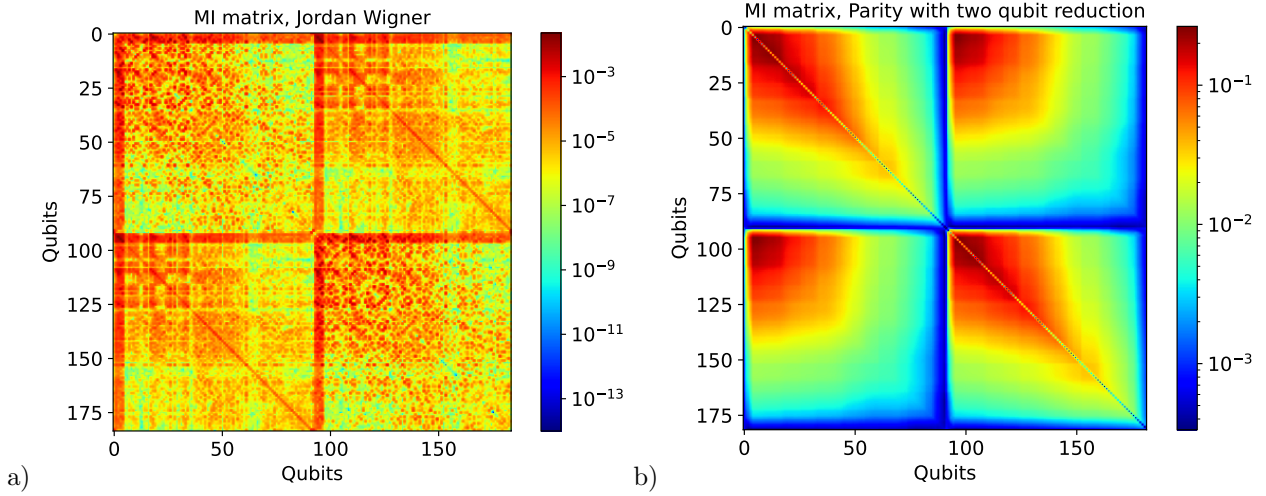
Figure 2: Mutual information for the water molecule at CISD level. Jordan Wigner mapping is reported on the left panel, parity mapping with the two-qubit reduction on the right panel. The starting wavefunction for the mutual information was the simulation for the water molecule(H_2O) as detailed in Tab. 1.

The results are summarized in Fig. 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 scaling method described by Eq. (29), while image b) shows the scaling of the direct tracing method, as explained in section 4.4.

Looking at image 1-a), we can infer that the scaling of $O(\chi^2 N)$ for the (29) 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 1-b), leading to the scaling of $O(\chi * N)$, as evaluated in section 4.4.1. This is because the shown simulation of 1 didn't make use of the particle-hole notation explained in appendix C specifically to show the behavior of the used Hash table.

This discussion leads us to believe that both Eq. (29) and algorithm 1 are valid methods to retrieve the partial trace and mutual information. However, while the quadratically scaling method of Eq. (29) can hardly handle up to approximately $\sim 10^5$ wavefunction components (1-a), the latter method, main-taining a linear computational cost as shown in 4.4.1, 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.

Wavefunction in different mappings

The first proof of concept involves encoding the fermionic wavefunction using different fermion-to-qubit mappings, as illustrated in Fig. 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 C and B.

By using Fig. 2, we notice that the Jordan-Wigner mapping (2-a) retains some chemical intuition, while the parity mapping (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[47]. 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.[48] 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)[49]. 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).
  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[21]. 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. (23)). 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 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 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[31;50;51] and quantum chemistry[52;53].

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.
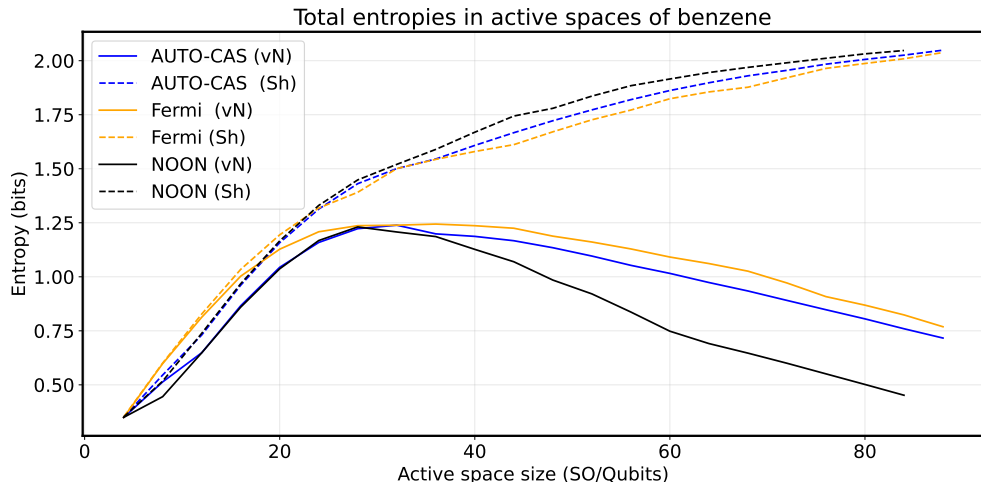
Figure 3: Total entropies as a function of the active space dimension and selection methods for a benzene molecule.The wavefunction analyzed is the CISD wavefunction of benzene with cc-pvdz basis. The Natural Orbitals for the natural orbitals occupation number (NOON) method were retrieved by an iterative 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.

7 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 Jordan-Wigner. This method enables the manipulation and analysis of large-scale quantum systems, significantly extend-ing 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.

7.1 Acknowledgments

Appendix A Mapping generalization

Even if the main focus of quantum information theory for quantum chemistry has always been within the Jordan-Wigner mapping, one might also be taken in account, for quantum computing reasons [54;31], 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:

$$\mathcal{H}(a_{q,j}^\dagger) \geqslant 1 \qquad (A.1)$$

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

Now, while the Jordan-Wigner [29] mapping, defined as in (10) can be easily proven to satisfy equality in (A.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. (11), we split each pairs of Majorana strings as follow

$$
\begin{aligned}
a_{q,\mathbf{i}}^\dagger \ket{\mathbf{0}} &= \prod_{l=1}^{N}(a_{q,l}^\dagger)^{\mathbf{i}_l} \ket{\mathbf{0}} = \\
&(a_{q,1}^\dagger)^{\mathbf{i}_1}\dots(a_{q,k}^\dagger)^{\mathbf{i}_k}\dots(a_{q,N}^\dagger)^{\mathbf{i}_N}\ket{\mathbf{0}} = \\
&\prod_{l=1}^{N}(\frac{1}{2}\gamma_q^{2l-1}-\frac{i}{2}\gamma_q^{2l})^{\mathbf{i}_l}\ket{\mathbf{0}} = \\
&\sum_{k=1}^{2^N} d_k \left(\bigotimes_{m=1}^{N}\sigma_{k,m}\right)\ket{\mathbf{0}}
\end{aligned}
\qquad (A.2)
$$

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 Jordan-Wigner encoding method, this expression

reduce to Eq.12 because the Hamming distance for each operator is equal to one. For a generic wavefunction

$$|\psi\rangle = \sum_{\mathbf{i}=\{0,1\}^N} c_{\mathbf{i}} a_{q,\mathbf{i}}^\dagger |\mathbf{0}\rangle \qquad (A.3)$$

we obtain

$$|\psi\rangle = \sum_{\mathbf{i}=\{0,1\}^N} c_{\mathbf{i}} \sum_{k=1}^{2^N} d_k \left(\bigotimes_{m=1}^N \sigma_{k,m} \right) |\mathbf{0}\rangle \qquad (A.4)$$

that become

$$|\psi\rangle = \sum_{j=1}^{4^N} \lambda_j \left(\bigotimes_{m=1}^N |\psi\rangle_{j,m} \right) \quad |\psi\rangle_{j,m} = \epsilon_{j,m} |0\rangle_m \qquad (A.5)$$

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 Jordan-Wigner encoding method, we obtain Eq.13, 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 4. 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

$$C_{\mathcal{K}}(|\psi\rangle) = \max_{\mathbf{i}\in|\psi\rangle} 2^{P_{\mathbf{i}}} \qquad (A.6)$$

where

$$P_{\mathbf{i}} = \sum_{l=1}^N \mathbf{i}_l \Theta(\mathcal{H}(a_{q,l}^\dagger) - 1) \qquad (A.7)$$

with

$$\Theta(x) = \begin{cases} 0 & \text{if x} = 0 \\ 1 & \text{if x} > 0 \end{cases} \qquad (A.8)$$

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

Appendix B   Efficient mapping permutation

Given that fermion-to-qubit mappings are a relatively active field of research[55;56;57;30;58], 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 Jordan-Wigner mapping in terms of the information it encodes.

However, it can be shown[59] 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.

$$U_{JW\to P} = CNOT_{N-1,N} \ldots CNOT_{i-1,i} \ldots CNOT_{0,1}$$
$$a_{P,i}^\dagger = U_{JW\to P} a_{JW,i}^\dagger U_{JW\to P}^\dagger$$
$$(B.1)$$

Considering that each piece in superposition of (13) for our qubit wavefunction is generated by $a_{q,\mathbf{i}}^\dagger$, we can rewrite

$$a_{P,\mathbf{i}}^\dagger |\mathbf{0}\rangle_q = U_{JW\to P} a_{JW,\mathbf{i}}^\dagger U_{JW\to P}^\dagger |\mathbf{0}\rangle =$$
$$= U_{JW\to P} a_{JW,\mathbf{i}}^\dagger |\mathbf{0}\rangle = U_{JW\to P} \left(a_{JW,\mathbf{i}}^\dagger |\mathbf{0}\rangle \right) \qquad (B.2)$$

Since the zero state is an eigenstate of the CNOT and thus $U_{JW\to P}^\dagger |\mathbf{0}\rangle = |\mathbf{0}\rangle$. At last, since $U_{JW\to P} \left(a_{JW,\mathbf{i}}^\dagger |\mathbf{0}\rangle \right)$ 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 Jordan Wigner mapping. Finally, we notice that since the only requirement used to establish (B.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 Jordan Wigner mapping, further details can be found in[59].

Appendix C   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

$$|\varnothing\rangle \rightarrow |\mathbf{0}\rangle \qquad (C.1)$$

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

$$|\mathbf{HF}\rangle \rightarrow |\mathbf{0}\rangle \qquad (C.2)$$

It follow that the operators of the fermionic modes $a_i^\dagger$, with $i = 1, \ldots, N$, are transformed as follow

$$a_i^\dagger = \begin{cases} a_i^\dagger & \text{if i} > n \\ b_i & \text{if i} \leqslant n \end{cases} \qquad (C.3)$$

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

$$\begin{aligned} \{b_i, b_j\} = \{a_k^\dagger, b_j^\dagger\} = \{a_k, b_j^\dagger\} = 0 \\ \{b_i, b_j^\dagger\} = \delta_{i,j} \mathbb{1} \end{aligned} \qquad (C.4)$$

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

$$a_l^\dagger a_m^\dagger a_n a_o |\mathbf{HF}\rangle = a_l^\dagger a_m^\dagger a_n a_o \prod_{i=1}^{n} a_i^\dagger |\varnothing\rangle \qquad (C.5)$$

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

$$a_l^\dagger a_m^\dagger a_n a_o |\mathbf{HF}\rangle \rightarrow a_{q,l}^\dagger a_{q,m}^\dagger a_{q,n} a_{q,o} |\mathbf{0}\rangle \qquad (C.6)$$

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.

References

[1] Tanya Zelevinsky, Artur F. Izmaylov, and Anastassia N. Alexandrova. Physical chemistry of quantum information science. *The Journal of Physical Chemistry A*, 127(49):10357–10359, 2023. PMID: 38093691.

[2] Richard P. Feynman. Simulating physics with computers. *International Journal of Theoretical Physics*, 21:467–488, 1982.

[3] Peter W Shor. Polynomial-time algorithms for prime factorization and discrete logarithms on a quantum computer. *SIAM Journal on Computing*, 26, 1997.

[4] Mari Carmen Bañuls, Rainer Blatt, Jacopo Catani, Alessio Celi, Juan Ignacio Cirac, Marcello Dalmonte, Leonardo Fallani, Karl Jansen, Maciej Lewenstein, Simone Montangero, Christine A. Muschik, Benni Reznik, Enrique Rico, Luca Tagliacozzo, Karel Van Acoleyen, Frank Verstraete, Uwe-Jens Wiese, Matthew Wingate, Jakub Zakrzewski, and Peter Zoller. Simulating lattice gauge theories within quantum technologies. *The European Physical Journal D*, 74:165, 8 2020.

[5] Hasham Hussain, Muhammad Bin Javaid, Faisal Shah Khan, Archismita Dalal, and Aeysha Khalique. Optimal control of traffic signals using quantum annealing. *Quantum Information Processing*, 19:312, 9 2020.

[6] 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.

[7] Szabo Attila and Ostlund Neil S. *Modern Quantum Chemistry: Introduction to Advanced Electronic Structure Theory*. Dover Publications, 2012.

[8] Alan Aspuru-Guzik, Anthony D. Dutoi, Peter J. Love, and Martin Head-Gordon. Simu-

lated quantum computation of molecular energies. *Science*, 309:1704–1707, 9 2005.

[9] Harper R. Grimsley, Sophia E. Economou, Edwin Barnes, and Nicholas J. Mayhall. An adaptive variational algorithm for exact molecular simulations on a quantum computer. *Nature Communications*, 10, 2019.

[10] Leonardo Ratini, Chiara Capecci, Francesco Benfenati, and Leonardo Guidoni. Wave function adapted hamiltonians for quantum computing. *Journal of Chemical Theory and Computation*, 18:899–909, 2022.

[11] Leonardo Ratini, Chiara Capecci, and Leonardo Guidoni. Optimization strategies in wahtor algorithm for quantum computing empirical ansatz: a comparative study. *Electronic Structure*, 5(4):045006, oct 2023.

[12] Daniel J. Egger, Chiara Capecci, Bibek Pokharel, Panagiotis Kl. Barkoutsos, Laurin E. Fischer, Leonardo Guidoni, and Ivano Tavernelli. A study of the pulse-based variational quantum eigensolver on cross-resonance based hardware. *Physical Review Research*, pages 1–13, 2023.

[13] Mark R. Hoffmann and Jack Simons. A unitary multiconfigurational coupled-cluster method: Theory and applications. *The Journal of Chemical Physics*, 88:993, 8 1998.

[14] Werner Kutzelnigg. Error analysis and improvements of coupled-cluster theory. *Theoretica Chimica Acta*, 80:349–386, 1991.

[15] Bridgette Cooper and Peter J. Knowles. Benchmark studies of variational, unitary and extended coupled cluster methods. *The Journal of Chemical Physics*, 133:234102, 12 2010.

[16] Francesco A. Evangelista. Alternative single-reference coupled cluster approaches for multireference problems: The simpler, the better. *The Journal of Chemical Physics*, 134:224102, 6 2011.

[17] James D. Whitfield, Jacob Biamonte, and Alan Aspuru-Guzik. Simulation of electronic structure hamiltonians using quantum computers. *Molecular Physics*, 109:735–750, 3 2011.

[18] G. Barcza, Ö. Legeza, K. H. Marti, and M. Reiher. Quantum-information analysis of electronic states of different molecular structures. *Physical Review A*, 83:012508, 1 2011.

[19] Katharina Boguslawski and Paweł Tecmer. Orbital entanglement in quantum chemistry. *International Journal of Quantum Chemistry*, 115:1289–1295, 10 2015.

[20] Lexin Ding, Sam Mardazad, Sreetama Das, Szilárd Szalay, Ulrich Schollwöck, Zoltán Zimborás, and Christian Schilling. Concept of orbital entanglement and correlation in quantum chemistry. *Journal of Chemical Theory and Computation*, 17:79–95, 2021.

[21] Christopher J. Stein and Markus Reiher. Automated selection of active orbital spaces. *Journal of Chemical Theory and Computation*, 12:1760–1771, 2016.

[22] Steven R. White. Density matrix formulation for quantum renormalization groups. *Phys. Rev. Lett.*, 69:2863–2866, 1992.

[23] Steven R. White. Density-matrix algorithms for quantum renormalization groups. *Phys. Rev. B*, 48:10345–10356, 1993.

[24] 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.

[25] U. Schollwöck. The density-matrix renormalization group. *Rev. Mod. Phys.*, 77:259–315, 2005.

[26] Ulrich 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.

[27] Katharina Boguslawski, Paweł Tecmer, Paul W. Ayers, Patrick Bultinck, Stijn De Baerdemacker, and Dimitri Van Neck. Efficient description of strongly correlated electrons with mean-field cost. *Physical Review B*, 89:201106, 5 2014.

[28] Moritz Bensberg and Markus Reiher. Corresponding active orbital spaces along chemical reaction paths. *The Journal of Physical Chemistry Letters*, 14(8):2112–2118, 2023. PMID: 36802629.

[29] P. Jordan and E. Wigner. Über das paulische Äquivalenzverbot. *Zeitschrift fur Physik*, 47:631–651, 9 1928.

[30] Aaron Miller, Zoltán Zimborás, Stefan Knecht, Sabrina Maniscalco, and Guillermo García-Pérez. Bonsai algorithm: Grow your own fermion-to-qubit mappings. *PRX Quantum*, 4:030314, 8 2023.

[31] Davide Materia, Leonardo Ratini, Celestino Angeli, and Leonardo Guidoni. Quantum information driven ansatz (qida): shallow-depth empirical quantum circuits from quantum chemistry. 9 2023.

[32] Sebastian Wouters and Dimitri Van Neck. The density matrix renormalization group for ab initio quantum chemistry. *The European Physical Journal D*, 68:272, 9 2014.

[33] Chr Møller and M S Plesset. Note on an approximation treatment for many-electron systems. *Physical Review*, 46, 1934.

[34] Jörg Rissler, Reinhard M. Noack, and Steven R. White. Measuring orbital interaction using quantum information theory. *Chemical Physics*, 323:519–531, 4 2006.

[35] Katharina Boguslawski, Paweł Tecmer, Gergely Barcza, Örs Legeza, and Markus Reiher. Orbital entanglement in bond-formation processes. *Journal of Chemical Theory and Computation*, 9:2959–2973, 7 2013.

[36] Frank Neese, Taras Petrenko, Dmitry Ganyushin, and Gottfried 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.

[37] Kantharuban Sivalingam, Martin Krupicka, Alexander A. Auer, and Frank Neese. Comparison of fully internally and strongly contracted multireference configuration interaction procedures. *The Journal of Chemical Physics*, 145(5):054104, 08 2016.

[38] Kerstin Andersson, Per-Ake Malmqvist, and Björn 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.

[39] 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.

[40] Michael A. Nielsen and Isaac L. Chuang. *Quantum Computation and Quantum Information*. Cambridge University Press, 2010.

[41] John von Neumann. Mathematische grundlagen der quantenmechanik. *Mathematische Grundlagen der Quantenmechanik*, 1996.

[42] Zhen Huang and Sabre Kais. Entanglement as measure of electron–electron correlation in quantum chemistry calculations. *Chemical Physics Letters*, 413:1–5, 9 2005.

[43] Thom 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.

[44] Rick A. Kendall, Thom H. Dunning, and Robert 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.

[45] L. Dagum and R. Menon. Openmp: an industry standard api for shared-memory programming. *IEEE Computational Science and Engineering*, 5:46–55, 1998.

[46] Qiming Sun, Timothy C Berkelbach, Nick S Blunt, George H Booth, Sheng Guo, Zhendong Li, Junzi Liu, James D McClain, Elvira R Sayfutyarova, Sandeep Sharma, Sebastian Wouters, and Garnet Kin-Lic Chan. Pyscf: the python-based simulations of chemistry framework. *WIREs Computational Molecular Science*, 8:e1340, 2018.

[47] Daniel S. Levine, Diptarka Hait, Norm M. Tubman, Susi Lehtola, K. Birgitta Whaley, and Martin 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.

[48] Jeppe Olsen. The casscf method: A perspective and commentary. *International Journal of Quantum Chemistry*, 111:3267–3272, 11 2011.

[49] Abhishek Khedkar and Michael Roemelt. Active space selection based on natural orbital occupation numbers from ¡i¿n¡/i¿ -electron valence perturbation theory. *Journal of Chemical Theory and Computation*, 15:3522–3536, 6 2019.

[50] Leonardo Ratini, Chiara Capecci, and Leonardo Guidoni. Natural orbitals and sparsity of quantum mutual information. *Journal of Chemical Theory and Computation*, 20(9):3535–3542, 2024.

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

[52] G. Berthier, A. Daoudi, and J.P. Flament. Natural orbitals in the cipsi algorithm. *Journal of Molecular Structure: THEOCHEM*, 166:81–86, 6 1988.

[53] 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.

[54] Zi-Jian Zhang, Thi Ha Kyaw, Jakob S Kottmann, Matthias Degroote, and Alan Aspuru-Guzik. Mutual information-assisted adaptive variational quantum eigensolver. *Quantum Science and Technology*, 6:035001, 7 2021.

[55] Sergey B. Bravyi and Alexei Yu Kitaev. Fermionic quantum computation. *Annals of Physics*, 298:210–226, 2002.

[56] Sergey Bravyi, Jay M. Gambetta, Antonio Mezzacapo, and Kristan Temme. Tapering off qubits to simulate fermionic hamiltonians. 1 2017.

[57] Zhang Jiang, Amir Kalev, Wojciech Mruczkiewicz, and Hartmut Neven. Optimal fermion-to-qubit mapping via ternary trees with applications to reduced quantum states learning. *Quantum*, 4:276, 6 2020.

[58] Aaron Miller, Adam Glos, and Zoltán Zimborás. Treespilation: Architecture- and state-optimised fermion-to-qubit mappings, 2024.

[59] Brent Harrison, Dylan Nelson, Daniel Adamiak, and James 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.

Acronyms

CAS Complete Active Space. 12, 13

CC Coupled Cluster. 3

CI Configuration Interaction. 3, 5, 12

CISD CI with singles and doubles. 10, 12, 13

DM Density Matrix. 5

DMRG Density Matrix Renormalization Group. 2, 3, 12

HF Hartree-Fock. 2, 3, 12, 13

MP2 Møller-Plesset of second order. 3

MPS Matrix Product State. 2, 5

NO Natural Orbitals. 12

NOON Natural Orbitals Occupation Number. 12

QC Quantum Computing. 2

SD Slater Determinants. 3, 5, 10

SO Spin Orbitals. 3

SparQ Sparse Quantum State Analysis. 2, 13, 14