

THE OPEN-SOURCE PYTHON PACKAGE QUANTUMSYMMETRY: THE WATER MOLECULE ON FOUR QUBITS

Dario Picozzi^{1} and Jonathan Tennyson¹*

¹*Department of Physics and Astronomy, University College London (UCL), Gower Street, London, WC1E 6BT, United Kingdom*

**picozzi.dario@gmail.com*

Keywords: quantum computing, molecular chemistry, point-group symmetries, complete active space, Python

Abstract

The symmetry-adapted encodings (SAEs) provide a useful tool to reduce the computational requirements of the simulation of molecular systems on quantum computers. This is achieved by harnessing the physical symmetries of the molecular system under consideration. They can also enable qubit-based complete active space approximation. The present work explores aspects of the symmetry-adapted encoding in the form of an open-source Python software, QuantumSymmetry, by considering a simple example of molecular system, that of the water molecule.

1 Introduction

Previous work has introduced the symmetry-adapted encodings (SAEs) [1] as a useful tool to reduce the computational requirements of the simulation of molecular systems on quantum computers. This is achieved by harnessing the physical symmetries of the molecular system under consideration, and specifically certain subgroups of the point-group (i.e. geometrical symmetries) and of number of up/down electron operators (spin along the z-axis being conserved under the purely non-relativistic approximation). The theory of symmetry-adapted encodings has been further expanded to perform qubit-based complete active space approximation: this is the subject of further work currently in the manuscript phase.

The present work explores aspects of the symmetry-adapted encoding in the form of an open-source Python software, QuantumSymmetry [2], by

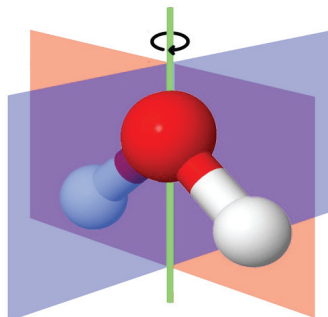


Fig. 1: The water molecule (H_2O) and its four point-group symmetries. Reproduced with permission from [1].

considering one simple example of a molecular system, that of the water molecule (H_2O). The theoretical aspects will only be touched on and the reader is invited to review the existing literature.

2 Symmetries of the water molecule

The water molecule in its optimal geometry has four point-group symmetries, which together form the point-group group C_{2v} : the identity element E (which is equivalent to not doing anything at all), the rotation by 180 degrees along the z-axis $C_2(z)$ (shown in Fig. 1 as a green line), the reflection across the xz -plane $\sigma(xz)$ (blue plane), and the reflection across the yz -plane $\sigma(yz)$ (red plane).

If a water molecule is rotated by 180 degrees around its principal axis its two hydrogen atoms exchange places, but the resulting geometry is indistinguishable from the one before the rotation was performed: in this sense, the 180-degree rotation around the principal axis $C_2(z)$ is a symmetry of the water molecule. The geometrical symmetries of the water molecule form what is known as its point group. Each point-group symmetry in C_{2v} repeated twice gives the identity (for example $C_2(z) \times C_2(z) = E$): C_{2v} is a Boolean group.

The conservation of the number of electrons in spin up N_\uparrow and spin down N_\downarrow furnish further symmetries. More in detail, the symmetries employed by the symmetry-adapted encodings are the parity (even/odd) of the number of spin up and down $P_\uparrow = (-1)^{N_\uparrow}$ and $P_\downarrow = (-1)^{N_\downarrow}$.

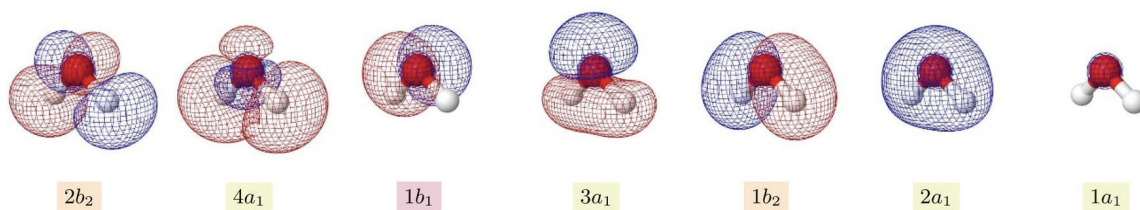


Fig. 2: The 7 water molecule molecular orbitals in a minimal basis labeled according to their point-group irreducible representation (in order of increasing energy from the right, following the little endian convention common in quantum computing). Reproduced with permission from [1].

The symmetries described above taken together allow for a reduction in qubit count of up to 4 qubits for the water molecule examples: 2 qubits are reduced due to the point-group symmetries and a further 2 qubits are reduced due to the parity conservation symmetries. In general, the number of qubits reduced from symmetries is between 2 and 5 depending on the level of symmetry of the molecule under consideration.

3 Jordan-Wigner Hamiltonian for the water molecule in the STO-3G basis

The Jordan-Wigner fermion-to-spin encodings maps electronic of a molecule states to qubit states on a quantum computer in the most natural way: every molecular orbital corresponds to two spin-orbitals (each corresponding to an electron with spin up or spin down), and each spin-orbital is mapped in one-to-one correspondence to a qubit. Slater determinants (electronic states that correspond to well-defined occupations of spin-orbitals) are thus sent to computational basis states, and superpositions of Slater determinants are mapped to the corresponding superpositions of computational basis states.

The molecular orbitals for the water molecule in a minimal basis (STO-3G basis) are shown in Fig 2.

The Jordan-Wigner encodings maps physical operators such as the second-quantised Hamiltonian (the operator associated with the energy of the electrons of the molecule) into a weighted sum of Pauli terms, i.e. tensor products of one-qubit Pauli operators X , Y and Z acting on the n qubits, where n is the number of spin-orbitals.

We are going to consider the water molecule in the configuration where the oxygen atom is positioned at $(0, 0.1167)$ and the two hydrogen atoms are positioned at $(0.7562, -0.4668)$ and $(-0.7562, -0.4668)$, with distances are measured in \AA .

The Jordan-Wigner Hamiltonian for the water molecule in the STO-3G basis has 1,086 Pauli terms acting on 14 qubits. The first few terms are of the form:

$$H = -46.41219 + 12.41370 Z_0 + 12.41370 Z_1 + 1.65714 Z_2 + 1.65714 Z_3 + \dots$$

where energy is measured in Ha.

The eigenvalue that corresponds to the ground state energy for this Hamiltonian is $E = -75.011754935519$ Ha

A circuit for a typical VQE [3] ansatz such as unitary coupled clusters with singles and double excitations (UCCSD) displays measures of complexity that make the circuit practically unfeasible on current quantum devices, e.g. a circuit depth (the number of gates in the longest path in the circuit) and CNOT gate count (the total number of CNOT gates in the circuit) of at least thousands of gates.

4 The symmetry-adapted Hamiltonian for the water molecule in the STO-3G basis

By harnessing the physical symmetries of the water molecule, the symmetry-adapted encoding is able to reduce the number of qubits necessary for the simulation of the water molecule from 14 qubits to 10 qubits. This is done by restricting the Hilbert space of qubits to one corresponding to a common eigensector of the symmetry operators under consideration.

For example, when the object of interest is the ground state of the water molecule, the Hartree-Fock state (the single Slater determinant that best approximates the ground state) is closed-shell and hence in the symmetry eigensector that is symmetric (eigenvalue equal to 1) with respect to each of the point-group operators $C_2(z)$, $\sigma(xz)$ and $\sigma(yz)$. The ground state is antisymmetric (eigenvalue equal to -1) with respect to the parity of spin up and down operators P_\uparrow and P_\downarrow , as there is an odd number (5) of electrons with spin up/down.

These constraints mean that of the 14 spin-orbital occupancies, 2 are made redundant by the point-group constraints (one of the point-group operators in this example does not add an independent constraint, as for example $C_2(z) = \sigma(xz) \times \sigma(yz)$, so that symmetry with

respect to both $\sigma(xz)$ and $\sigma(yz)$ already implies symmetry with respect to $C_2(z)$: we just need consider the generators of the point group), and a further 2 are made redundant by the parity operators.

Thus 4 of the original 14 qubits can be discarded, and the occupancy of the corresponding spin-orbitals can be deduced from the remaining 10 qubits. Once the choice of redundant qubits has been made, the symmetry-adapted encoding provides a way to construct the qubit operators corresponding to physical operators such as the Hamiltonian.

In particular the Hamiltonian for water in the STO-3G basis under the symmetry-adapted encoding can be obtained by use of the following code in QuantumSymmetry:

```
reduced_hamiltonian(atom = 'O O O 0.1167; H O
0.7562 -0.4668; H O -0.7562 -0.4668',
basis = 'sto-3g',
charge = 0,
spin = 0
)
```

The user can set the optional argument `output_format = 'qiskit'` to obtain the symmetry-adapted Hamiltonian as a Qiskit [4] Pauli sum operator object or to `'openfermion'` to obtain an OpenFermion [5] qubit operator object (defaults to OpenFermion).

The output is the sum of 1,035 Pauli terms acting on 10 qubits. The first few Pauli terms are of the form:

$$H = -46.19215 - 12.41370 Z_8 Z_6 Z_5 Z_3 Z_2 Z_0 - 12.41370 Z_8 Z_7 Z_5 Z_4 Z_2 Z_1 + 2.74250 Z_5 + 1.65714 Z_0 + \dots$$

The eigenvalue that corresponds to the ground state energy of the symmetry-adapted Hamiltonian is $E = -75.011754935519$ Ha, the same as the Jordan-Wigner Hamiltonian to the order of 10^{-12} Ha, showing the general feature that symmetry-adapted encoding with exact symmetries are an exact procedure that does not affect the value of measurable energies.

While the symmetry-adapted Hamiltonian gives rise to a UCCSD circuit whose circuit depth and CNOT count reduced by a factor of 5 with respect to the original Jordan-Wigner Hamiltonian, the complexity of such a circuit is still prohibitive for current quantum devices.

5 The water molecule in active space CAS(4, 4)

The active space approximation [6] classifies molecular orbitals into three categories: frozen core

orbitals, active space orbitals and virtual orbitals. The frozen core orbitals are assumed to be fully occupied, the virtual ones are assumed to be unoccupied, and a certain number of electrons are allowed to be excited from and to the active space orbitals only. In the literature it is common to use the notation $CAS(n, m)$ to specify the number of active electrons n , and the number of active molecular orbitals m .

The formalism of the symmetry-adapted encodings can be employed to perform qubit-based complete active space approximation (the subject of further work currently in the manuscript phase). Furthermore the formalism that has been developed can be compounded with the existing formalism for point-group and parity symmetries, leading to a further reduction in complexity.

For the water molecule example, a CAS(4, 4) approximation would result in 8 active-space spin orbitals. This can be encoded in just 4 qubits if the active orbitals are taken to be $3a_1$, $1b_1$, $4a_1$ and $2b_2$ (respectively orbital 4, 5, 6 and 7), by employing the point-group and parity symmetries.

The user can do this by setting the optional argument `CAS = (4, 4)` when calling the `reduced_hamiltonian` function.

The resulting Hamiltonian has just 62 terms acting on 4 qubits. The first few terms are of the form:

$$H = -72.73209 + 0.85365 Z_3 Z_0 + 0.60520 Z_3 Z_1 + 0.60520 Z_3 Z_2 + 0.57203 Z_2 Z_0 + 0.57203 Z_1 Z_0 + \dots$$

It is important to note that the CAS approximation will introduce an error in the ground state energy. In this case the ground state energy corresponding to the CAS symmetry-adapted Hamiltonian is $E = -74.969911454158$ Ha. However when the CASCI energy for the same active space is evaluated with a traditional quantum chemistry software such as PySCF, the corresponding energy is found to be $E = -74.969911454158$ Ha, giving again an agreement in the order of 10^{-12} Ha.

The resulting UCCSD circuit is reduced in complexity by orders of magnitude, with circuit depth of 134 gates and CNOT-count of 64 gates.

However, it was found that the VQE on the linear TwoLocal circuit [7] on 4 qubits with Ry gates and CNOTs and two repeated blocks converges to a value of the ground state energy that is closer to the exact one, with an even lower complexity, with circuit depth of 16 and CNOT-count of 12, allowing for practical simulation of larger molecules such as the water molecule on existing quantum computers.

6 Active space selection

The importance of active space optimization in quantum chemistry computations was noted in [8]. With respect to the water molecule example, the authors note that a choice of active space for CAS(4, 4) that captures a larger portion of the correlation energy is given by $1b_2$, $3a_1$, $4a_1$ and $2b_2$ (respectively orbital 3, 4, 6 and 7).

The user is able to manually specify which orbitals are selected to form the active space by setting the optional argument `active_mo = [3,4,6,7]`. This results in the qubit Hamiltonian made up of 132 terms on 5 qubits:

$$H = -72.95527 + 0.46991 Z_2 Z_0 + 0.34536 Z_2 Z_1 + 0.31342 Z_4 Z_3 + 0.30563 Z_4 Z_3 Z_2 Z_0 + \dots$$

This time the qubit reduction due to exact symmetries is of 3 qubits (as opposed to 4 qubits). This is because the active-space orbitals selected are all in either the A_1 or the B_2 irreducible point group representation, effectively reducing the symmetry from the C_{2v} point group to the smaller C_2 group.

The ground state energy associated with this Hamiltonian has energy -74.994776500138 Ha, again showing agreement to 10^{-12} Ha with the energy calculated with a traditional quantum chemistry package.

7 Larger basis set

The qubit-based active space approximation allows us to consider active spaces in a larger basis than the minimal basis at the cost of more classical preprocessing but no extra cost in terms of quantum computing resources. The output Hamiltonian for a larger basis set with CAS(4, 4), namely Dunning's correlation consistent basis sets cc-pVDZ [9], has the same number of Pauli terms (132 terms on 5 qubits).

$$H = -74.30120 + 0.47453 Z_4 Z_2 Z_0 + 0.47453 Z_3 Z_2 Z_0 + 0.43360 Z_0 + 0.43360 Z_2 Z_1 Z_0 + \dots$$

The exact ground state energy for this example is -76.03016669691 Ha, with agreement to 10^{-11} Ha with the energy calculated with a traditional quantum chemistry package.

8 Excited states

By selecting a different target irreducible representation (setting the optional argument `irrep = 'B1'`) QuantumSymmetry is able to construct a qubit Hamiltonian whose lowest eigenvalue is the first excited state of the water molecule, with an energy of -75.546986405341 Ha for the cc-pVDZ basis set with CAS(4, 4) and the same active space selection as in

the last two sections. This is made up of 121 terms on 5 qubits:

$$H = -74.30120 I + 0.47453 Z_4 Z_2 Z_0 + -0.47453 Z_3 Z_2 Z_0 + 0.43360 Z_0 + -0.43360 Z_2 Z_1 Z_0 + 0.19500 Z_3 Z_1 + 0.19500 Z_4 Z_2 + \dots$$

9 Conclusion

Examples of applications of symmetry-adapted encodings have been presented in relation to a simple molecular system, that of the water molecule. In particular, we have shown that although the problem is too complex for the framework of current quantum devices when approached through the customary Jordan-Wigner encoding, it simplifies considerably and becomes tractable when the exact procedure of symmetry-adapted encodings is employed in combination with the complete active space approximation, allowing for molecules to be simulated on a systems of a few (e.g. 4-5) qubits and for the measure of related circuit complexity with applications such as the VQE reduced by various orders of magnitude.

10 Acknowledgements

DP's research is supported by an industrial CASE (iCASE) studentship, funded by the Engineering and Physical Sciences Research Council (EPSRC) [EP/T517793/1], in collaboration with University College London and Rahko Ltd.

11 References

- [1] Picozzi, D. and Tennyson, J. (2023) 'Symmetry-adapted encodings for qubit number reduction by point-group and other boolean symmetries', *Quantum Science and Technology*, 8(3), p. 035026. doi:10.1088/2058-9565/acd86c.
- [2] Picozzi, D. (2023) *QuantumSymmetry*, doi:10.5281/zenodo.7724697.
- [3] Peruzzo, A. *et al.* (2014) 'A variational eigenvalue solver on a photonic quantum processor', *Nature Communications*, 5(1). doi:10.1038/ncomms5213.
- [4] Qiskit contributors (2023), *Qiskit: An Open-source Framework for Quantum Computing*. doi:10.5281/zenodo.2573505.
- [5] McClean, J.R. (2017) *et al.*, *OpenFermion: The electronic structure package for quantum computers ...*, *arXiv*. Available at: <https://arxiv.org/abs/1710.07629>.

- [6] Jensen, F. (2017) *Introduction to computational chemistry*. New Jersey: John Wiley & Sons Inc.
- [7] *TwoLocal - Qiskit documentation*. Available at: <https://qiskit.org/documentation/stubs/qiskit.circuit.library.TwoLocal.html>.
- [8] de Gracia Triviño, J.A., Delcey, M.G. and Wendin, G. (2023) ‘Complete active space methods for NISQ devices: The importance of canonical orbital optimization for accuracy and Noise Resilience’, *Journal of Chemical Theory and Computation*, 19(10), pp. 2863–2872. doi:10.1021/acs.jctc.3c00123.
- [9] Dunning, T.H. (1989) ‘Gaussian basis sets for use in correlated molecular calculations. i. the atoms boron through neon and hydrogen’, *The Journal of Chemical Physics*, 90(2), pp. 1007–1023. doi:10.1063/1.456153.