Accuracy assessment of quantum computing algorithms for chemistry: a numerical study of variational Ansätze across a database of problems

IBM authors

IBM Quantum, IBM Research Almaden, 650 Harry Road, San Jose, CA 95120, USA

Virginia Tech authors

Department of Chemistry, Virginia Tech, 480 Davidson Hall, Blacksburg, VA 24061, USA

I. INTRODUCTION

Recent years have witnessed remarkable research in the simulation of many-body quantum systems by quantum computing algorithms. A prominent application targeted by quantum algorithms is the electronic structure (ES) problem, namely solving for the ground or low-lying eigenstates of the electronic Schrödinger equation for atoms, molecules, and materials. Research in this direction has delivered promising results in the calculation of potential energy curves, ground- and excited-state energies and correlation functions for various instances of the ES problem.

Notwithstanding these rapid developments, the behavior of quantum computing algorithms for the ES problem remains to be understood and established, and such a situation is due to several factors. For example: different research groups demonstrate existing or novel algorithms on different instances of the ES problem, which makes comparative algorithm analysis difficult; the design and implementation of novel quantum computing algorithms is often coupled to their demonstration on contemporary quantum hardware, which restricts the domain of potential applications to very small numbers of electrons and orbitals; the limitations of classical simulators have similarly resulted in most simulations reported to date being restricted to very small instances of the ES problem, so that the observed algorithmic behavior is not guaranteed to generalize to larger or more complex instances; even when quantum algorithms have known limitations, their numeric assessment is often limited by the immaturity of contemporary quantum computational platforms, which makes difficult to gauge the severity of such limitations.

Overcoming these limiting factors can contribute to the understanding of the nature, performance and limitations of quantum algorithms for the ES problem, and bring progress in their design and assessment. Useful steps towards this goal are: to establish a database of problems that different researchers can use to test quantum algorithms; to propose a set of metrics to quantify algorithmic performance, and support an objective comparison of algorithmic behavior; to establish a database of results that different researchers can consult to compare and contrast quantum algorithms.

The proposed problems should be sufficiently simple, to be demonstrated on contemporary quantum computing platforms, and non-trivial, to ensure meaningful assessments of algorithmic behavior. Moreover, the assessment of algorithmic behavior should decouple classical simulators of quantum devices and actual quantum devices, so that theoreticians can explore the properties of a methodology assuming an ideal quantum device, and experimentalists can independently test a set of well-defined quantum circuits motivated by the solution of instances of the ES problem.

In this work, we present a set of instances of the ES problem, and perform an assessment of Ansätze used in variational ground-state quantum simulations, using classical simulators of quantum hardware. In the assessment of quantum algorithms, we will use the difference between the computed and exact ground-state energy as an indicator of algorithmic quality, along with the expectation values of quantities signaling the breaking of Hamiltonian symmetries, a conceptually important but often overlooked aspect of variational quantum simulations. We will use a variety of indicators of algorithmic cost, related to the optimization of variational parameters (number of variational parameters, cost of energy evaluation) and to the quantum circuits leading to the calculation of energies and properties (number of qubits, gates and measurements, circuit depth).

As illustrative applications, we employ the database presented here to test a new variational Ansatz termed *cascade* against existing methods, to compare open- and closed-shell implementations of the unitary coupled-cluster with singles and doubles [1, 2], and to compare two different strategies to encode fermionic degrees of freedom to qubits.

The remainder of this work is organized as follows: in Section II we describe the proposed database, and the algorithms tested in the present work; in Section III we present the results of this study; conclusions and perspective prosecutions of this work are describe in Section IV, and an appendix contains additional computational details.

II. METHODS

In this section, we present the problems constituting the database studied in this work, and the metrics used to establish algorithm performance and computational cost.

A. Target problems

In this work, we focus on a set of four molecular species, namely BH, HF, BeH_2, and H_2O. The BH, HF, BeH and OH bondlengths R vary between 0.5 and 3.5, 4.5, 4.5, and 3.3 Angstrom respectively. These molecular species, proposed in previous works, allow to test quantum algorithms along the breaking of one (BH and HF) or two (BeH_2, and H_2O) covalent bonds, and in presence of dominant dynamic (small R) or static (large R) electronic correlation.

For each species, a restricted closed-shell Hartree-Fock calculation is performed, as specified in Appendix D. The core 1s orbital is frozen and the Born-Oppenheimer Hamiltonian

$$\hat{H} = E_0 + \sum_{\substack{pr \\ \sigma}} h_{pr} \hat{c}^\dagger_{p\sigma} \hat{c}_{r\sigma} + \sum_{\substack{prqs \\ \sigma\tau}} \frac{(pr|qs)}{2} \hat{c}^\dagger_{p\sigma} \hat{c}^\dagger_{q\tau} \hat{c}_{s\tau} \hat{c}_{r\sigma} \tag{1}$$

is computed, along with the number, spin-z, and total spin operators

$$
\begin{aligned}
\hat{N} &= N_f + \sum_{p\sigma} \hat{c}^\dagger_{p\sigma} \hat{c}_{p\sigma} \quad, \\
\hat{S}_z &= \sum_{p\sigma} (-1)^\sigma \hat{c}^\dagger_{p\sigma} \hat{c}_{p\sigma} \quad, \\
\hat{S}^2 &= \sum_{pq} \hat{c}^\dagger_{p\downarrow} \hat{c}_{p\uparrow} \hat{c}^\dagger_{q\uparrow} \hat{c}_{q\downarrow} + \hat{S}_z(\hat{S}_z - 1) \quad,
\end{aligned}
\tag{2}
$$

where $N_f = 2$ is the number of frozen electrons. In this study we focused on a minimal STO-6G basis, to ensure a number of orbitals and qubits within the reach of contemporary classical simulators of quantum circuits.

B. Qubit encodings

In the majority of quantum algorithms for contemporary devices, the Hamiltonian Eq. (1) and the auxiliary operators Eq. (2) are mapped onto qubit operators, and more specifically onto linear combinations

$$\hat{H} = \sum_{i=0}^{n_p-1} c_i \, \sigma_{\mathbf{v}_i \mathbf{w}_i} \quad, \quad \sigma_{\mathbf{v}\,\mathbf{w}} = \bigotimes_{k=0}^{n_q-1} \sigma_{v_k w_k} \quad, \tag{3}$$

of Pauli operators acting on a certain number n_q of qubits. In Equation (3) $\sigma_{00,01,10,11} = \mathbb{1}, X, Z, Y$ is a single-qubit Pauli operator, c_i a real-valued coefficient, and the summation ranges over n_p terms.

In this study, we employ both second-quantization and first-quantization encodings. The former establish a one-to-one mapping between the Fock space of electrons in M spatial orbitals and the Hilbert space of $n_q = 2M$ qubits, and the latter between the Hilbert space spanned by a set of K space- and spin-adapted configuration state functions (CSFs) and the Hilbert space of $n_q = \lceil \log_2 K \rceil$ qubits.

First-quantization encodings ensure access to automatically symmetry-adapted many-electron wavefunctions, and require less qubits than their second-quantization counterparts. However, they are less straightforward to implement, and lead to a representation of the Hamiltonian where n_p scales exponentially with n_q. Additional technical details are provided in Appendix B.

Second-quantization encodings give rise to a compact representation of the Hamiltonian and of auxiliary operators as linear combinations of a polynomial number of Pauli operators, but allow the possibility that Hamiltonian symmetries are broken, for multiple reasons discussed in Section II D. To alleviate the latter issue, we combine the parity mapping with the two-qubit and tapering techniques. The former removes two qubits to ensure conservation of the parity operators $(-1)^{\hat{N}_\uparrow}$ and $(-1)^{\hat{N}_\downarrow}$, and the latter identifies a subgroup of the symmetry group isomorphic to $\mathbb{Z}_2^{\times k}$ and removes k qubits to ensure the simulated wavefunctions lie in a target irreducible representation of the identified subgroup. The species studied here have symmetry groups ($C_{\infty v}$ for BH and HF, $D_{\infty h}$ for BeH_2 and C_{2v} for H_2O) isomorphic to $\mathbb{Z}_2 \times \mathbb{Z}_2$, so that the combined use of two-qubit reduction and tapering leads to $n_q = 2M - 4$ qubits.

C. Target variational Ansätze

In this study, we elected to explore the behavior of the variational quantum eigensolver (VQE) algorithm, due to its widespread adoption in the community. This method defines a set of Ansatz states approximating the ground state of a target Hamiltonian, of the form

$$|\Psi(\boldsymbol{\theta})\rangle = \hat{U}(\boldsymbol{\theta})|0\rangle^{\otimes n_q} \quad , \tag{4}$$

where $\boldsymbol{\theta} \in [0, 2\pi)^{n_\theta}$ is an array of angles that define a parametrized quantum circuit $\hat{U}(\boldsymbol{\theta})$, applied to a register of n_q qubits prepared in the standard initial state $|0\rangle^{\otimes n_q}$. The best approximation to the ground state in the set of Ansatz states is found by minimizing the energy

$$E_{\text{VQE}}(\boldsymbol{\theta}) = \langle\Psi(\boldsymbol{\theta})|\hat{H}|\Psi(\boldsymbol{\theta})\rangle \tag{5}$$

as a function of the parameters $\boldsymbol{\theta}$ using a classical optimization algorithm. Once the parameters are optimized, auxiliary observables can be measured, yielding results

$$X_{\text{VQE}}(\boldsymbol{\theta}) = \langle\Psi(\boldsymbol{\theta})|\hat{X}|\Psi(\boldsymbol{\theta})\rangle \quad . \tag{6}$$

This algorithmic workflow, termed variational quantum eigensolver (VQE) in the quantum simulation literature, is a heuristic technique for ground-state approximation. Its accuracy and computational cost, as discussed in Section II D, are determined by the details of the circuit $\hat{U}(\boldsymbol{\theta})$. In the remainder of this Section, we will present the Ansätze assessed in the present work.

1. Unitary coupled cluster with singles and doubles

The quantum unitary coupled cluster (q-UCC) method is based on the exponential Ansatz, i.e., the exact wave function is written as

$$|\Psi_{\text{gs}}\rangle = e^{\hat{T}-\hat{T}^\dagger}|\Phi_0\rangle \tag{7}$$

where Φ_0 is an independent particle function (here, the Hartree-Fock state) and $\hat{T}$ is a cluster operator which, at the singles and doubles level (q-UCCSD) is truncated to

$$\begin{aligned}
\hat{T} &= \hat{T}_1 + \hat{T}_2 \quad , \\
\hat{T}_1 &= \sum_{\substack{ai \\ \sigma}} t_{i\sigma}^{a\sigma} \hat{c}_{a\sigma}^\dagger \hat{c}_{i\sigma} \quad , \\
\hat{T}_2 &= \sum_{\substack{abij \\ \sigma\tau}} t_{i\sigma\ j\tau}^{a\sigma\ b\tau} \hat{c}_{a\sigma}^\dagger \hat{c}_{b\tau}^\dagger \hat{c}_{j\tau} \hat{c}_{i\sigma} \quad ,
\end{aligned} \tag{8}$$

where t_i^a and t_{ij}^{ab} are a set of unknown cluster coefficients. We have adopted the convention that i, j, k, l and a, b, c, d refer to occupied and unoccupied orbitals in the reference configuration Φ_0 respectively. It should be noted that the form Eq. (8), presented in previous literature along with the corresponding quantum circuit, corresponds to an unrestricted Ansatz, which is not guaranteed to yield an eigenfunction of total spin.

If a closed shell spin-adapted formalism is used, then, following Paldus and Scuseria et al, we need to have

$$t_{i\alpha}^{a\alpha} = t_{i\beta}^{a\beta} = t_i^a \quad , \tag{9}$$

since the two other possible spin combinations are vanishing. Only 6 of the 16 possible spin combinations for the two-electron amplitudes are non-zero, namely

$$\begin{aligned}
t_{i\alpha\ j\alpha}^{a\alpha\ b\alpha} &= t_{i\beta\ j\beta}^{a\beta\ b\beta} = \hat{t}_{ij}^{ab} \\
t_{i\alpha\ j\beta}^{a\alpha\ b\beta} &= t_{i\beta\ j\alpha}^{a\beta\ b\alpha} = \tilde{t}_{ij}^{ab} \\
t_{i\alpha\ j\beta}^{a\beta\ b\alpha} &= t_{i\beta\ j\alpha}^{a\alpha\ b\beta} = \overline{t}_{ij}^{ab}
\end{aligned} \tag{10}$$

Using the relations

$$\hat{t}^{ab}_{ij} = \tilde{t}^{ab}_{ij} + \overline{t}^{ab}_{ij} \quad , \quad \tilde{t}^{ab}_{ij} = -\overline{t}^{ab}_{ji} = -\overline{t}^{ab}_{ij} = \tilde{t}^{ba}_{ji} \quad , \tag{11}$$

one is left with only one set of independent two-electron amplitudes $\tilde{t}^{ab}_{ij}$ and $(ai) \leq (bj)$. In addition to reducing the number of variational parameters, the use of a closed shell spin-adapted formalism ensures that the cluster operator commutes with the total spin operator, $[\hat{T}, \hat{S}^2] = 0$. Its exponential, however, cannot be implemented exactly at polynomial cost on a quantum computer. Instead, as discussed in Appendix A, a Trotter approximation is required, which in turn leads to spin symmetry breaking.

2. Hardware-efficient R_y

3. Cascade

D. Metrics of algorithm performance

A natural metric to assess the accuracy of a variational simulation is the difference between the computed and exact ground-state energy,

$$\Delta E = E_{\text{VQE}}(\boldsymbol{\theta}) - E_{\text{FCI}} \quad . \tag{12}$$

The ES Hamiltonian has several symmetries, i.e. operators $\hat{S}$ such that $[\hat{H}, \hat{S}] = 0$. In presence of Hamiltonian symmetries, the search for Hamiltonian eigenfunctions has to be restricted to eigenspaces of Hamiltonian symmetries with known eigenvalues. Relevant Hamiltonian symmetries are the auxiliary operators Eq. (2) and, for the molecules studied in the present work, the ground state lies in the eigenspaces of $\hat{S}_z$ and $\hat{S}^2$ with eigenvalue zero, and of $\hat{N}$ with eigenvalues 6 for BH and BeH$_2$ and 10 for HF and H$_2$O.

In general, a variational quantum simulation does not preserve Hamiltonian symmetries. To assess whether a variational simulation leads to symmetry breaking, and to quantify its extent, we compute the deviations

$$\begin{aligned} \Delta N &= N_{\text{VQE}}(\boldsymbol{\theta}) - N_{\text{FCI}} \quad , \\ \Delta S_z &= S_{z,\text{VQE}}(\boldsymbol{\theta}) - S_{z,\text{FCI}} \quad , \\ \Delta S^2 &= S^2_{\text{VQE}}(\boldsymbol{\theta}) - S^2_{\text{FCI}} \quad , \end{aligned} \tag{13}$$

between the computed and exact ground-state particle number, spin-z, and total spin.

In addition to the accuracy of a variational simulation, it is also important to quantify its computational cost. To this end we employ a tuple of values, namely: the number of qubits n_q, the number of gates n_g and the depth d of the variational circuit, the number of Pauli operators in the Hamiltonian n_p, and the number of variational parameters $n_{\boldsymbol{\theta}}$.
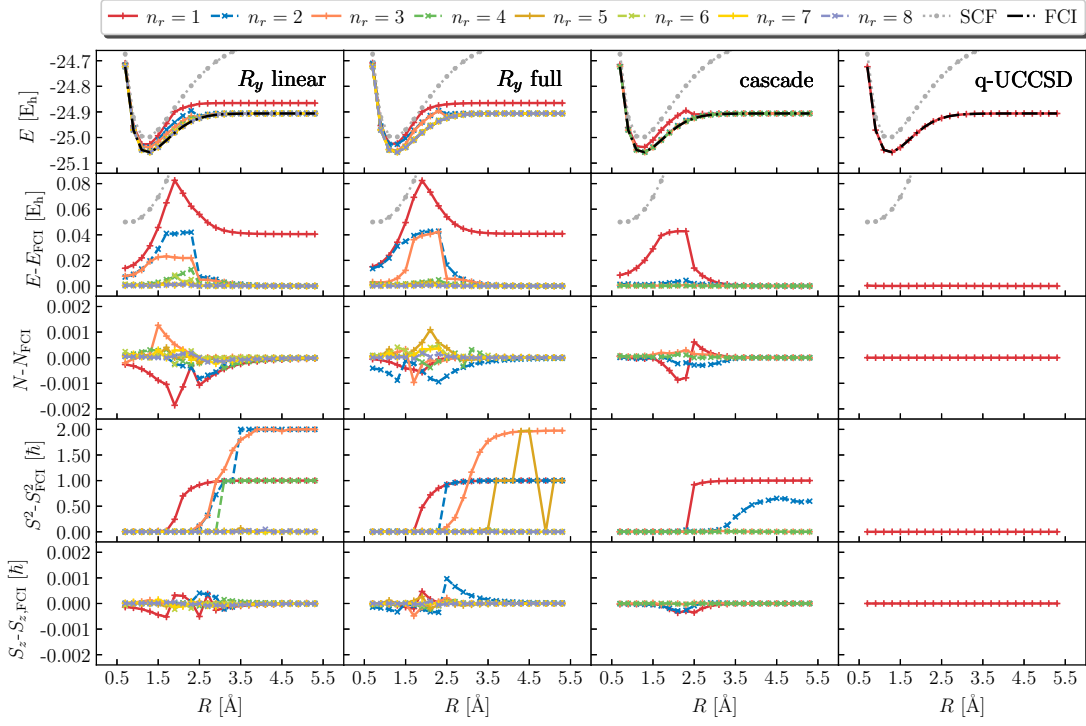
FIG. 1. Total energy, deviation between computed and exact total energy, electron number, total spin, and spin-z (top to bottom) using the R_y (with linear and full connectivity), cascade, and q-UCCSD Ansätze (left to right), for the BH molecule at STO-6G level. **smooth out Ry full spin**
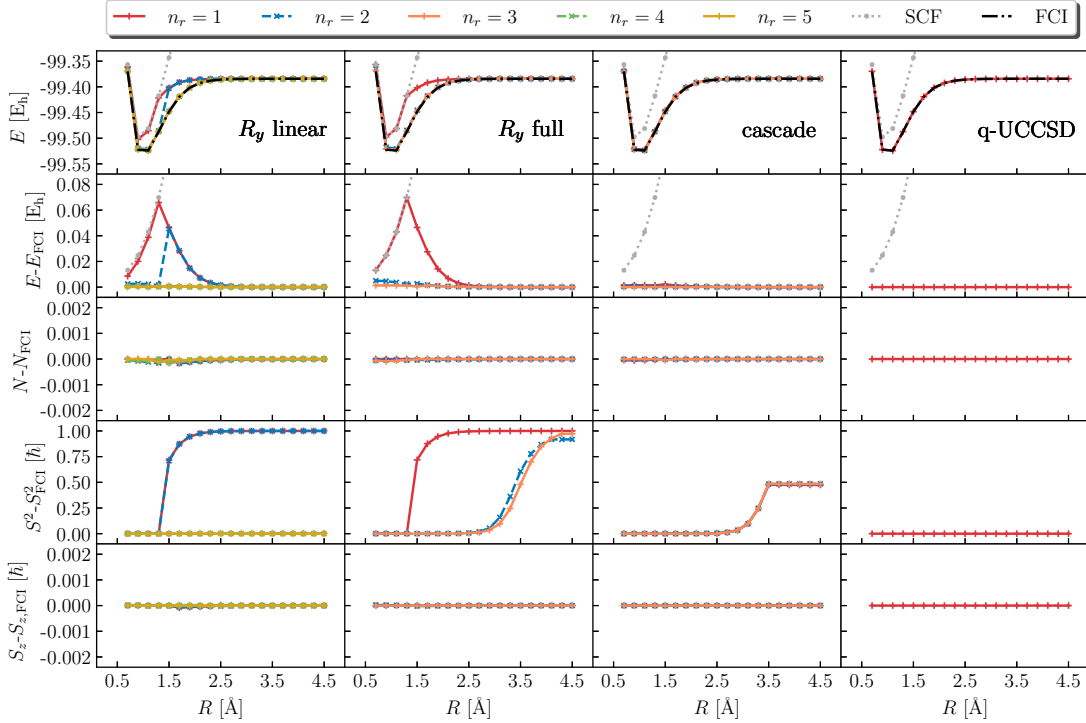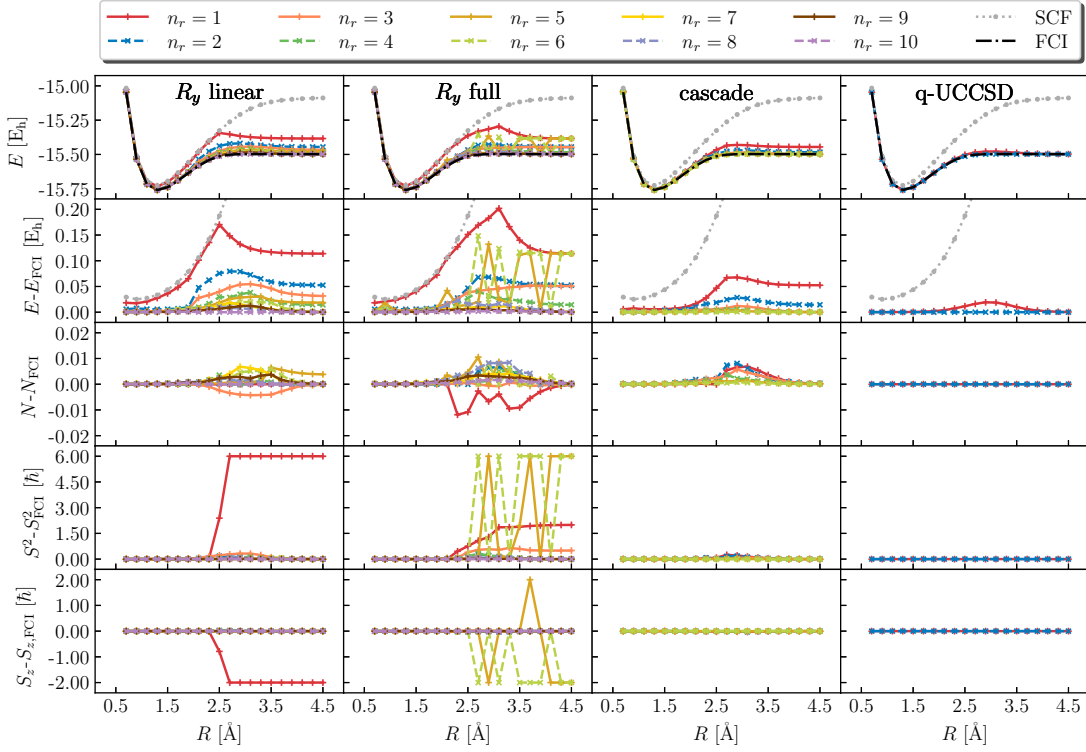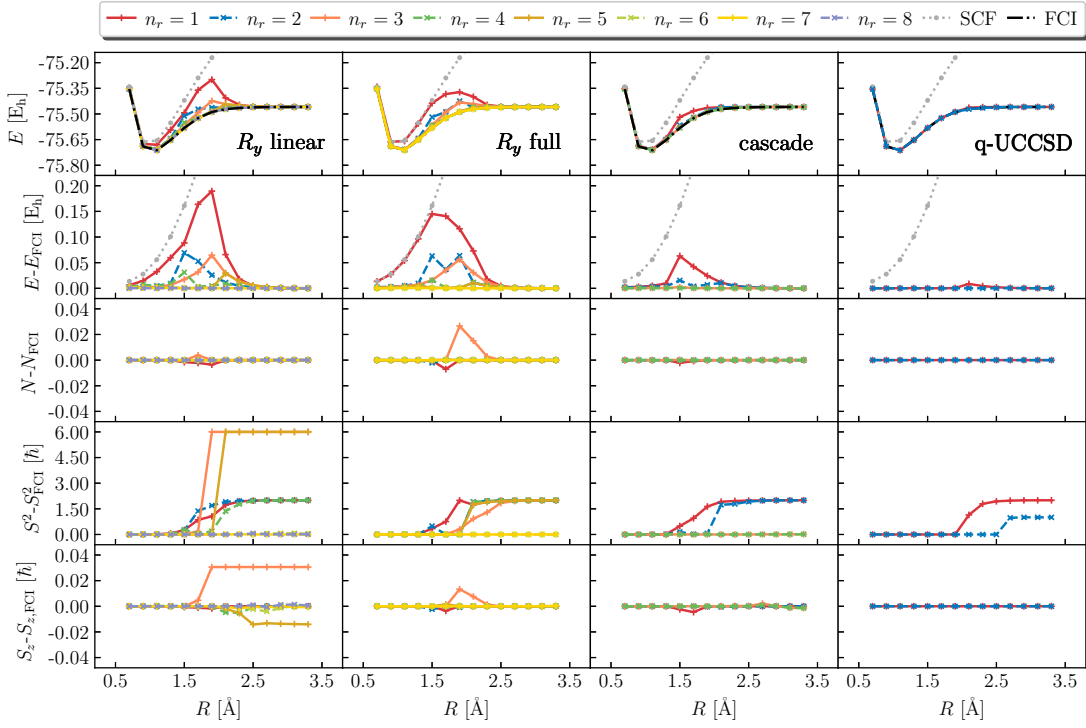

FIG. 2. Total energy, deviation between computed and exact total energy, electron number, total spin, and spin-z (top to bottom) using the R_y (with linear and full connectivity), cascade, and q-UCCSD Ansätze (left to right), for the HF molecule at STO-6G level.

FIG. 3. Total energy, deviation between computed and exact total energy, electron number, total spin, and spin-z (top to bottom) using the R_y (with linear and full connectivity), cascade, and q-UCCSD Ansätze (left to right), for the BeH_2 molecule at STO-6G level. **smooth out Ry full spin and $n_r = 9$ energy for Ry linear**

III. RESULTS

IV. CONCLUSIONS AND OUTLOOKS

ACKNOWLEDGMENTS

Appendix A: Closed-shell q-UCCSD

Appendix B: First quantization encoding

Appendix C: Projection after variation

Appendix D: Classical preprocessing

[1] G. E. Scuseria, A. C. Scheiner, T. J. Lee, J. E. Rice, and H. F. Schaefer III, The Journal of chemical physics **86**, 2881 (1987).

[2] P. G. Szalay and J. Gauss, The Journal of chemical physics **107**, 9028 (1997).

FIG. 4. Total energy, deviation between computed and exact total energy, electron number, total spin, and spin-z (top to bottom) using the R_y (with linear and full connectivity), cascade, and q-UCCSD Ansätze (left to right), for the H_2O molecule at STO-6G level.
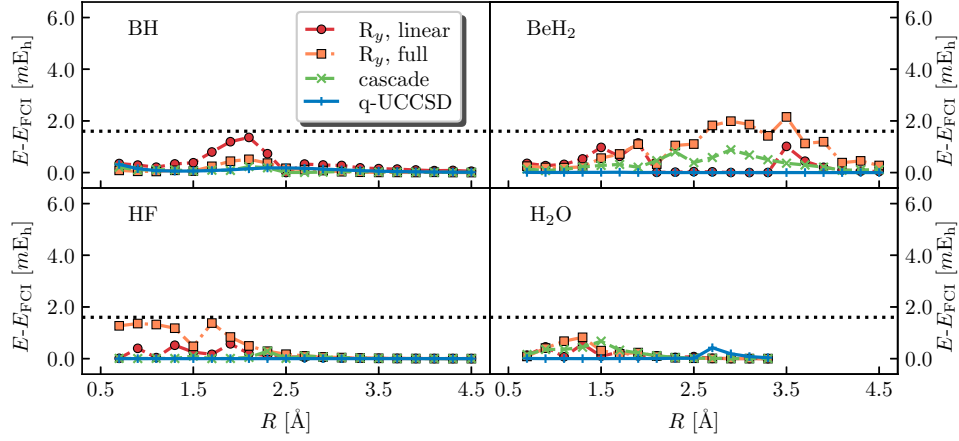

FIG. 5. Deviation between computed and exact total energy for the BH, HF, BeH$_2$ and H$_2$O molecules at STO-6G level (counterclockwise) using the R_y (with linear and full connectivity), cascade, and q-UCCSD Ansätze (red circles, orange squares, green crosses, and blue markers) with the highest computed number of repetitions. **look into BeH2 energies for Ry full**
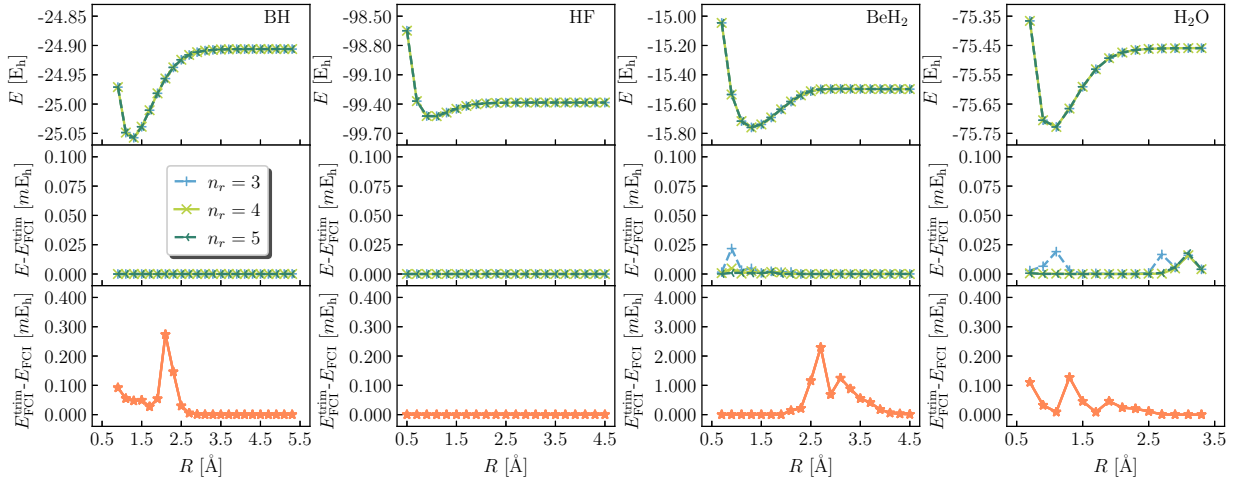
FIG. 6. Total energy, deviation between computed and exact total energy using a first-quantization encoding and a trimming procedure for qubit reduction, and energy error from the trimming procedure (top to bottom) for the BH, HF, BeH$_2$ and H$_2$O molecules at STO-6G level (left to right) using the cascade Ansatz.
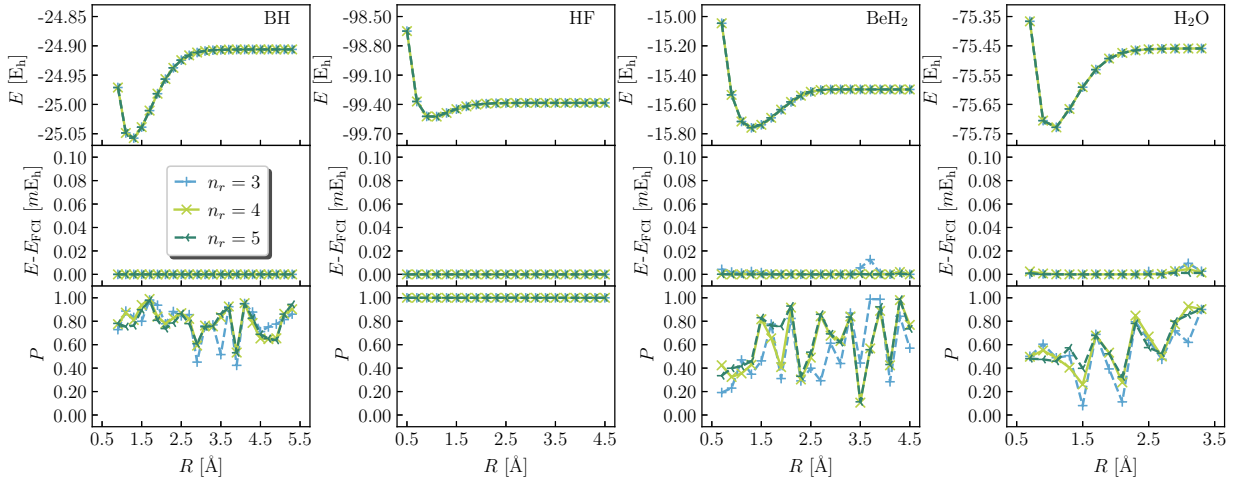

FIG. 7. Total energy, deviation between computed and exact energy, and squared-norm of the unphysical component of the wavefunction (top to bottom) for the BH, HF, BeH$_2$ and H$_2$O molecules at STO-6G level (left to right) using a first-quantization encoding with padding and a cascade Ansatz optimized with the variation-after-projection scheme.
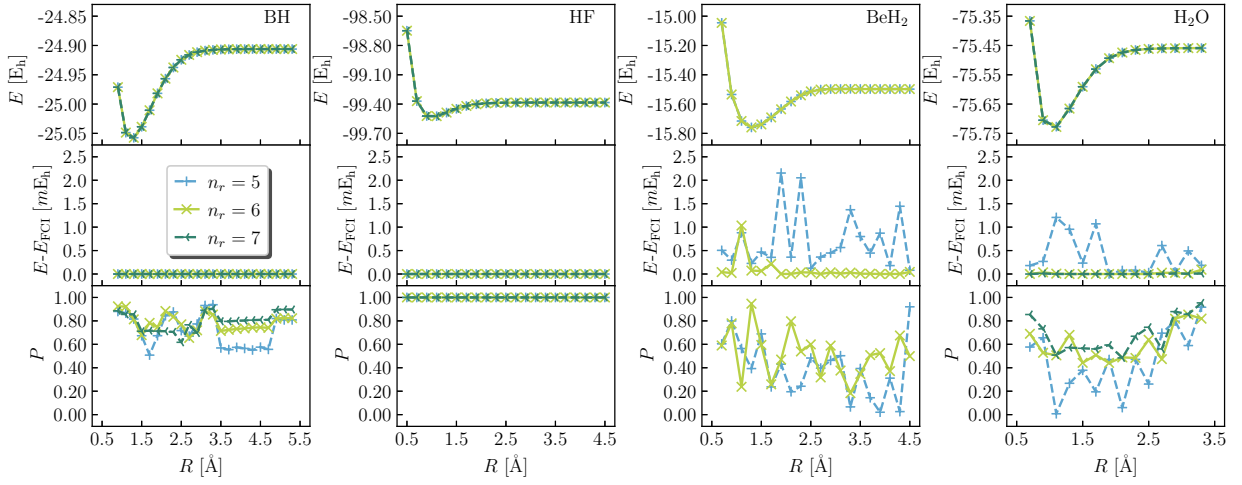
FIG. 8. Total energy, deviation between computed and exact energy, and squared-norm of the unphysical component of the wavefunction (top to bottom) for the BH, HF, BeH$_2$ and H$_2$O molecules at STO-6G level (left to right) using a first-quantization encoding with padding and an R$_y$ Ansatz with linear connectivity optimized with the variation-after-projection scheme. **BeH2 $n_r = 7$ is incomplete, smooth out H2O with depth $n_r = 5$ and BeH2 with depth $n_r = 5, 6$**
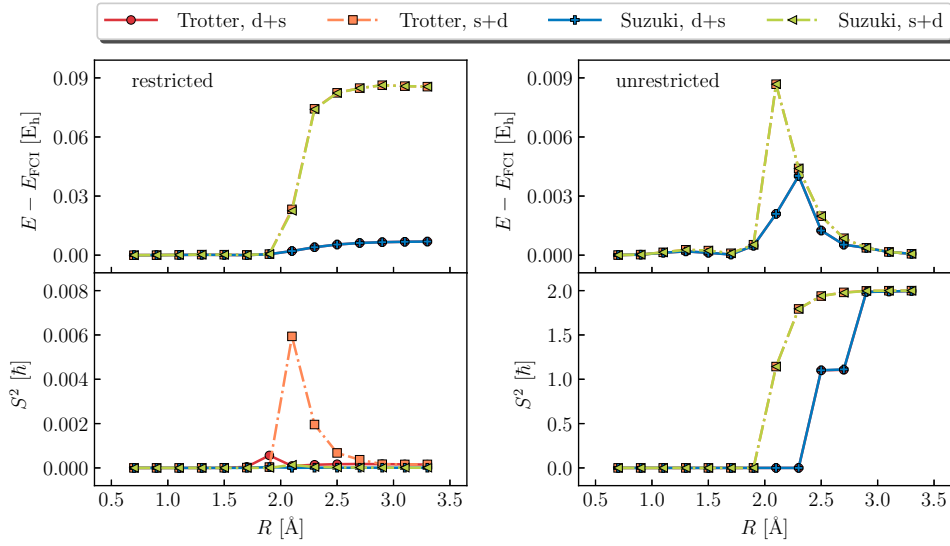

FIG. 9. Deviation between computed and exact energy (top) and total spin (bottom), for the restricted closed-shell (left) and unrestricted (right) qUCCSD Ansätze with $n_r = 1$ repetitions, for the H$_2$O molecule at STO-6G level, using Trotter and Suzuki approximations (warm, cold colors) and with singles followed by doubles and doubles followed by singles (light, dark colors).
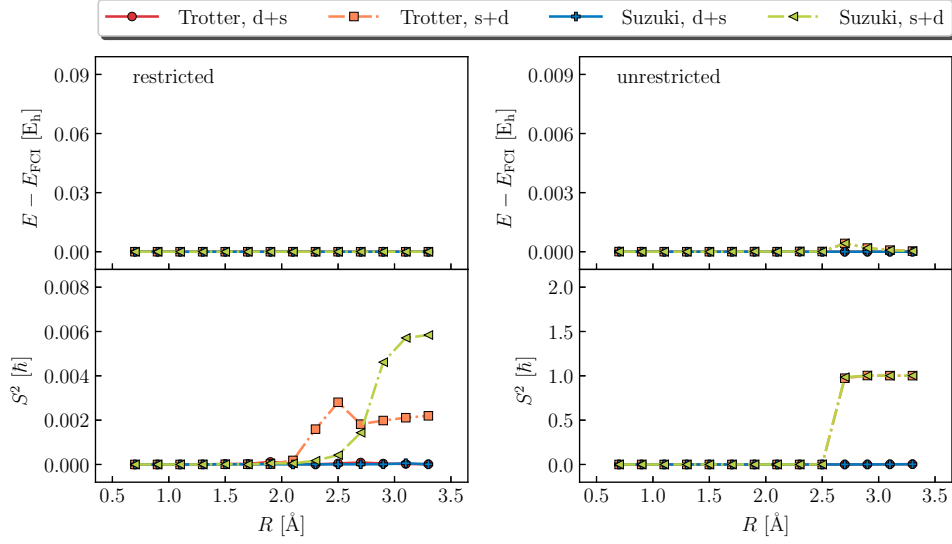
FIG. 10. Deviation between computed and exact energy (top) and total spin (bottom), for the restricted closed-shell (left) and unrestricted (right) qUCCSD Ansätze with $n_r = 2$ repetitions, for the H_2O molecule at STO-6G level, using Trotter and Suzuki approximations (warm, cold colors) and with singles followed by doubles and doubles followed by singles (light, dark colors).
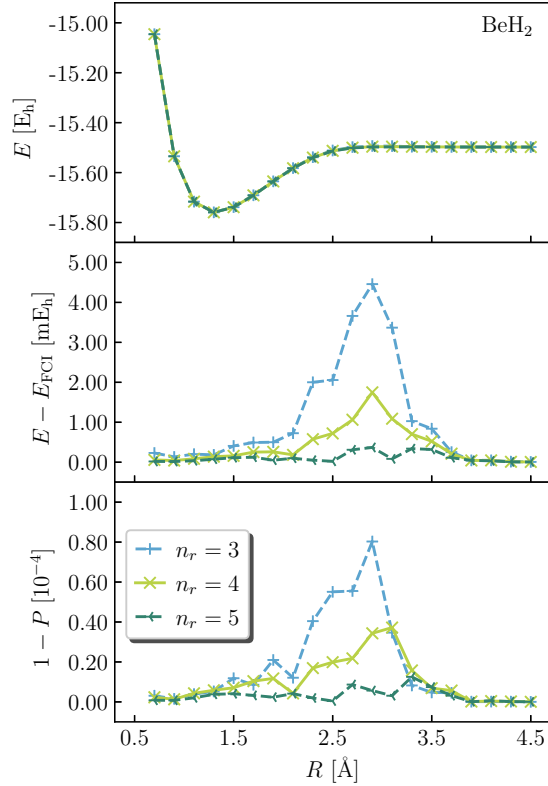

FIG. 11. Total energy, deviation between computed and exact energy, and squared-norm of the unphysical component of the wavefunction (top to bottom) for the BeH_2 molecule at STO-6G level, using a first-quantization encoding with padding and a cascade Ansatz optimized with the projection-after-variation scheme.
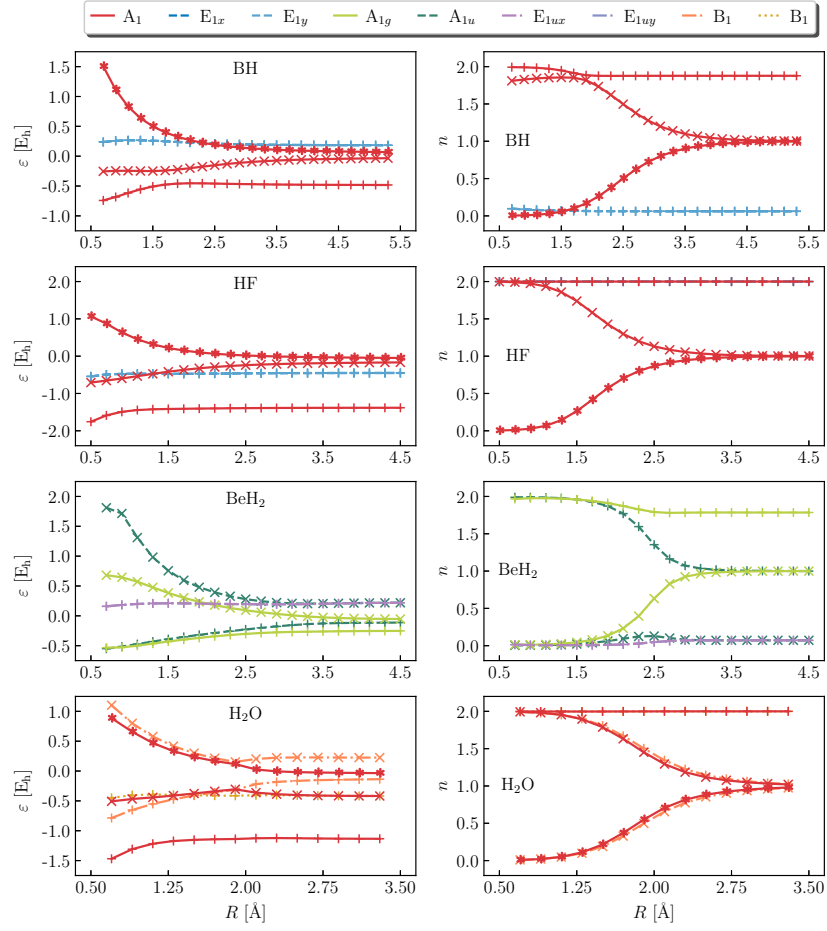
FIG. 12. Eigenvalues of the Fock operator in Hartree units (left, ε) and occupation numbers of the exact ground-state one-body density matrix (right, n) for BH, HF, BeH$_2$ and H$_2$O (top to bottom). Colors denote irreducible representations of the molecular point-group symmetry; plus, cross, and asterisk markers denote orbitals in ascending order of energy within a specific irreducible representation; the energy and occupation number of the core 1s orbital are omitted.