Ground-state-energy calculation for the water molecule on a superconducting quantum processor

Michael A. Jones,* Harish J. Vallury, and Lloyd C.L. Hollenberg
School of Physics, University of Melbourne, Parkville 3010, Australia

(Received 2 February 2024; revised 11 April 2024; accepted 13 May 2024; published 7 June 2024)

The accurate computation of properties of large molecular systems is classically infeasible and is one of the applications in which it is hoped that quantum computers will demonstrate an advantage over classical devices. However, due to the limitations of present-day quantum hardware, variational-hybrid algorithms introduced to tackle these problems struggle to meet the accuracy and precision requirements of chemical applications. Here, we apply the quantum computed moments (QCM) approach combined with a variety of noise-mitigation techniques to an eight qubit and spin-orbital representation of the water molecule (H_2O). A noise-stable improvement on the variational result for a four-excitation trial state (circuit depth 25, 22 CZ gates) was obtained, with the ground-state energy computed to be within 1.0 ± 0.8 mHa of exact diagonalization in the 14 spin-orbital basis. Thus, the QCM approach, despite an increased number of measurements and noisy quantum hardware (CZ error rates c.0.3% corresponding to expected error rates on the trial-state circuit of order 7%), is able to determine the ground-state energy of a nontrivial molecular system at the required accuracy (c.0.05%). To the best of our knowledge, these results are the largest calculations performed on a physical quantum computer to date to achieve millihartree agreement with exact diagonalization, and a promising indicator of how such hybrid approaches might scale to problems of interest in the low-error and fault-tolerant regimes as quantum computers develop.

DOI: 10.1103/PhysRevApplied.21.064017

I. INTRODUCTION

While quantum computing promises to revolutionize computational quantum chemistry in the long term (i.e., using millions of error-corrected qubits [1,2]), there is also considerable interest in whether near-term quantum devices can provide an advantage for chemistry applications. In this noisy intermediate-scale quantum [3] regime, the mainstay approach to computing quantities such as the ground-state energy, E_0, of molecular systems defined by a Hamiltonian, $\mathcal{H}$, are hybrid quantum-classical algorithms [4,5] based on the variational principle:

$$\langle \mathcal{H} \rangle = \langle \Psi_{\text{trial}} | \mathcal{H} | \Psi_{\text{trial}} \rangle \geq E_0. \tag{1}$$

The expectation is that, as the problem size scales, the preparation and measurement of the (parameterized) trial state, $|\Psi_{\text{trial}}\rangle$, on the quantum device will be less expensive than for the classical counterpart. Variational algorithms have been successfully applied to relatively small (≤ 6 qubit) systems [4,6–31] and/or restricted trial states (e.g., Hartree-Fock [22,32] or pair-correlated [33,34] states), with recent work [35,36] extending to higher qubit counts, though not yet achieving the targeted accuracy level. The key issue preventing scale up of the approach is the effect of device noise on the trial-state preparation, which in the context of required accuracy rapidly overwhelms the computation, pushing the variational upper bound away from the true ground-state energy.

A promising improvement on the variational approach proposed recently is to use a noise robust ground-state energy estimate from Lanczos expansion theory [37–40] based on higher-order Hamiltonian moments, $\langle \mathcal{H}^p \rangle$. The quantum computed moments (QCM [41]) approach has been seen to improve on the variational energy estimate in the presence of noise [42] and has been applied to hydrogen atom chains for restricted trial states up to six qubits [22]. This work seeks to extend the application of the QCM approach in both the number of qubits and trial-state complexity by computation of the ground-state energy of the water molecule (H_2O) using a unitary coupled cluster doubles (UCCD [43]) based ansatz defined over eight qubits and spin orbitals on a superconducting quantum processor through the workflow presented in Fig. 1. The trial circuits used encode each spin orbital into a separate qubit and implement four and six double excitations requiring 22 and 70 CZ gates, respectively (Fig. 2). Such an encoding is preferable to assigning one spatial orbital (two spin orbitals) per qubit, as it allows for more flexibility in trial-state design. On the other hand, encoding spatial orbitals

*Corresponding author: jonesm3@student.unimelb.edu.au

allows for much more compact trial states, even for the same number of qubits.

Here, we show that through the use of a carefully designed trial circuit, noise-mitigation techniques and the QCM approach, an error from exact diagonalization in the minimal STO-3G basis (14 spin orbitals) of 1–2 mHa is achieved for the four-excitation trial-state circuit. To the best of our knowledge, this is one of the largest quantum computed electronic structure calculations (in terms of the number of entangled qubits) that achieves millihartree accuracy (relative to FCI). Additionally, the computations are performed in a gate-based framework and do not employ techniques such as pulse optimization or long-term device monitoring, which could be implemented to further improve the results.

A. Quantum chemistry

Given a fixed nuclear geometry, the electronic structure Hamiltonian of a molecular system can be written in second quantization as

$$\mathcal{H} = \sum_{jk} h_{jk} a_j^\dagger a_k + \sum_{jklm} g_{jklm} a_j^\dagger a_k^\dagger a_l a_m, \tag{2}$$

where $a_j^\dagger$ (a_j) are the creation (annihilation) operators for an electron in spin orbital j, and h_{jk} (g_{jklm}) are the classically computed one- (two-) body electronic integrals defined by the chosen basis of spin orbitals [Figs. 1(a) and 1(b)]. Implicit in this definition is the discretization of space into a basis set of three-dimensional electronic wave functions, known as orbitals. Each orbital is composed of two spin orbitals (corresponding to the spin degree of freedom of a single electron) and this discretization of space introduces a level of approximation to the problem, with the error vanishing as the number of spin orbitals approaches infinity. Given the Hamiltonian, the problem is then to estimate the ground-state energy to high accuracy. The ultimate goal of quantum chemistry methods is to achieve "chemical accuracy," which is defined to be a result within 1.6 mHa of the infinite basis limit *or* within the same margin of accurate experimental methods. A slightly less demanding target is to achieve "FCI accuracy," where a result is within 1.6 mHa of full configuration interaction (FCI, equivalent to matrix diagonalization) for a given basis size. FCI accuracy has been achieved on noisy quantum devices for various small systems using up to five–six qubits [7,9,12,18,20–22,28,30,31,36]. Finally, "trial-state accuracy," can be defined as when the measured energy is within 1.6 mHa of the exact trial-state energy [15,32]. The results presented here

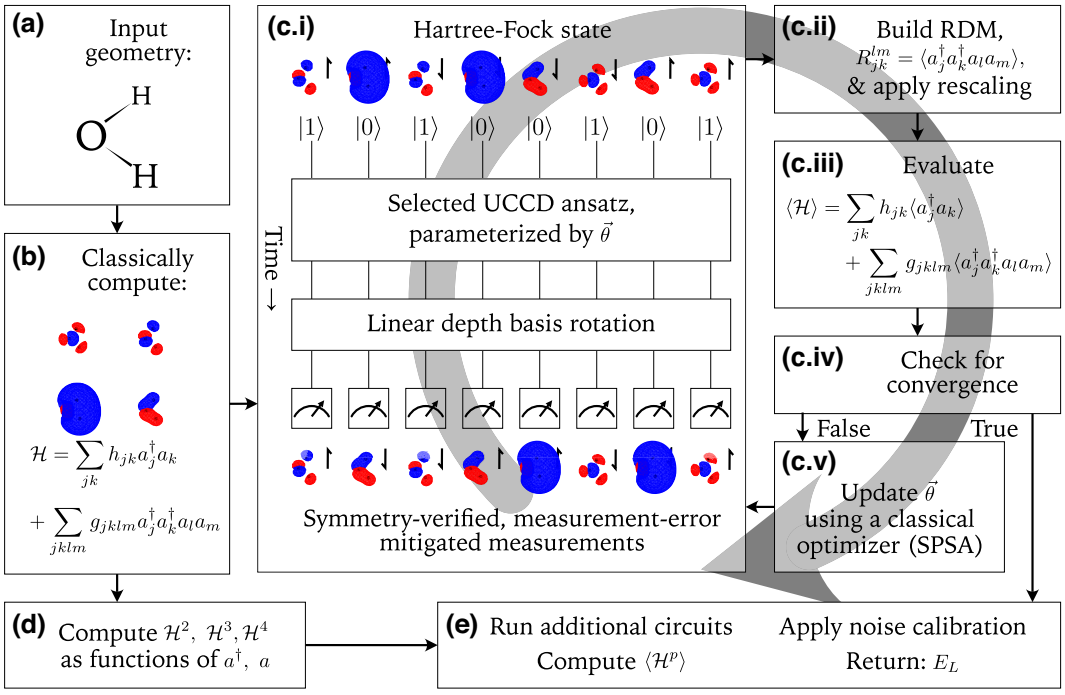

FIG. 1. Flow chart of the overall approach employed in this work. (a) Definition of the molecular system. (b) Calculation of the molecular orbitals and electronic Hamiltonian. (c) Hybrid minimization loop to approximate the ground state on the quantum device. (d) Powers of the Hamiltonian can be evaluated independently of (i.e., in parallel with) the minimization loop. (e) Calculation of the Hamiltonian moments, $\langle \mathcal{H}^p \rangle$, and hence the Lanczos cluster expansion-corrected energy estimate, E_L. For additional detail on the above steps see Sec. IV and the Appendices.

achieve FCI accuracy, though the FCI computation is performed over 14 spin orbitals while the quantum computation is performed over eight spin orbitals. Of the remaining six orbitals, two core orbitals are neglected completely and the four virtual orbitals are incorporated via the QCM approach.

B. The variational quantum eigensolver method

Expressing the electronic structure problem as a second-quantized Hamiltonian allows for a direct mapping between spin-orbital occupation states and qubit computational basis states, i.e., the Jordan-Wigner encoding [45]. In this way each spin orbital corresponds to one qubit and the problem is to construct a state, $|\Psi_{\text{trial}}\rangle = |\Psi(\vec{\theta})\rangle$, on the quantum computer and identify the values of $\vec{\theta}$ for which the energy is minimized [Fig. 1(c)]. The possibility of quantum advantage comes from the fact that, unlike classical computers, which require memory scaling exponentially with the number of spin orbitals, a quantum computer requires only a number of qubits linear in the number of spin orbitals to represent an arbitrary trial state. While completely general trial states will require exponentially deep circuits, it is possible to construct reasonable, physically motivated trial states in polynomial depth such as via the (Trotterized) UCC operator as discussed in Sec. IV and Appendix A. Once the parameterized quantum circuit is defined, the energy can be measured efficiently and the parameters optimized via a classical minimization loop [4]. From Eq. (1), this minimum is an upper bound to the true ground-state energy and, assuming no hardware noise, that the trial circuit is expressive enough, and the optimization has been performed well, it will be a tight upper bound.

C. The quantum computed moments method

The theory underpinning the energy correction used in the QCM method [41] is related to quantum subspace expansion, quantum Krylov methods and other quantum moments-based methods [9,24,46–62], that effectively generate a reduced Hamiltonian in a subspace that (ideally) includes the ground state. The implementation of the QCM method used here employs an analytic form for the corrected ground-state energy, obtained from Lanczos cluster expansion theory [37], to fourth order in the Hamiltonian moments, $\langle \mathcal{H}^p \rangle$ [38,39]:

$$E_L = c_1 - \frac{c_2^2}{c_3^2 - c_2 c_4}\left(\sqrt{3c_3^2 - 2c_2c_4} - c_3\right), \quad (3)$$

where the cumulants are given by

$$c_p = \langle \mathcal{H}^p \rangle - \sum_{j=0}^{p-2} \binom{p-1}{j} c_{j+1} \langle \mathcal{H}^{p-1-j} \rangle. \quad (4)$$

Here, expectation values are taken with respect to the optimized trial state. In addition to avoiding the explicit construction and diagonalization of the reduced Hamiltonian, Eq. (3) has been seen to be surprisingly robust against noise [42] in the computation of the moments compared to evaluation of $\langle \mathcal{H} \rangle$ alone. Another possible application is the extension of the QCM method to the computation of Green's functions as demonstrated recently [63]. In the context of chemical computations, the QCM method has been applied to electronic structure problems using single Slater-determinant states [22]. Here we employ multideterminant trial states, which require much deeper trial circuits and a range of error-mitigation techniques.

II. RESULTS

Calculations were performed following the overall procedure outlined in Fig. 1. The parameters, $\hat{\theta}$, in the Trotterized UCCD trial circuit were optimized using a classical

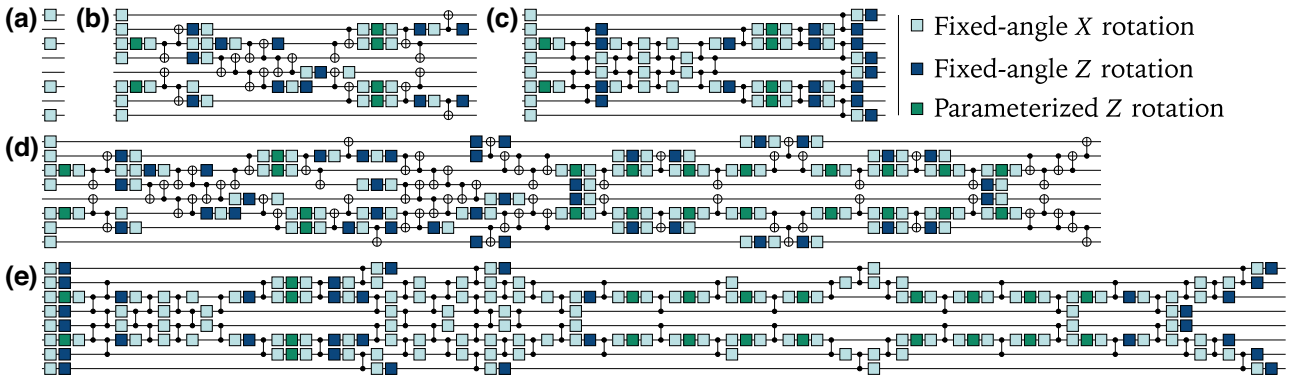

FIG. 2. Structure of the trial-state circuits used in this work [44]. (a) Minimal circuit to prepare the Hartree-Fock state. Four-excitation trial-state circuit compiled to the base gateset of (b) *ibmq_kolkata* and (c) *ibm_torino* each with depth $D = 25$. Six-excitation trial-state circuit compiled to the base gateset of (d) *ibmq_kolkata* and (e) *ibm_torino* with depths $D = 74$ and $D = 87$, respectively. For each circuit, there exist a set of parameters such that the Hartree-Fock state is produced. These circuits are followed by linear depth basis rotations (see Appendix B), then computational basis measurements.

state-vector simulator; the problem of optimizing the trial-state structure and parameters in the presence of noise is left to future work. The final computations were performed on the *ibmq_kolkata* and *ibm_torino* devices using 10^5 and 2.5×10^4 shots per basis, respectively. Additional method details can be found in Sec. IV and the Appendices.

Two different trial-state depths were used, the state preparation circuits in Figs. 2(b) and 2(c) have a depth of 25 (22 CNOT or CZ gates) and implement four excitations from the truncated UCC ansatz (see Appendix A) while the circuits in Figs. 2(d) and 2(e) have depths 74 and 87 (72 CNOT and 70 CZ gates, respectively) and implement six excitations. After performing the demonstration, a slightly more simplified form of circuit Fig. 2(d) was found (70 CNOTS, 73 depth). It is expected that the improved circuit will have minimal impact on the final energy estimates. These depths and gate counts include only the state-preparation component of the circuit since different measurement bases increase these values by different amounts. The results for both energy estimates $\langle \mathcal{H} \rangle$ and E_L and for both trial states are presented in Fig. 3 and numerical data is given in Table I. As expected the results from *ibm_torino* outperform the results from *ibmq_kolkata* since the former is a newer device with lower reported error rates, this is true even though the *ibmq_kolkata* results use 4 times as many shots as the *ibm_torino* results. A key observation regarding the results presented is that the moments of the Hartree-Fock state are classically tractable (and are required for the reference-state calibration technique used for noise mitigation, see Appendix C 4). Therefore, if the quantum computer is going to provide any utility in the computation, it must give expectation values that are (at least) competitive with this reference-state energy (black dashed lines in Fig. 3). It is also worthwhile to note that the energy computed from Eq. (3) is nonvariational, so the FCI energy is not a lower bound, which can lead to "negative errors."

A. Four-excitation trial state

For the four-excitation trial state, all energy estimates ($\langle \mathcal{H} \rangle$ and E_L on both devices) have lower error than the reference state. $\langle \mathcal{H} \rangle$ in particular, is separated from the corresponding reference-state estimate by more than two standard deviations, indicating that the benefits of implementing the quantum circuit outweigh the cost (in accuracy) imposed by device noise. For E_L, the separation between noise-mitigated and reference-state energies (in absolute error from FCI) is less than half a standard deviation for *ibmq_kolkata* and about one standard deviation for *ibm_torino*. The reduced resolution can be largely attributed to the high accuracy of the moments method applied to the reference state, which has an error only 0.5 mHa larger than the state-vector simulation. Given that the target accuracy is 1.6 mHa, the obtained standard deviations of $\leq$1.2 mHa are acceptable and it is unnecessary to attempt to reduce this further. Instead, it may be of interest to repeat the demonstration for a system that is not well modeled by Hartree-Fock methods to determine if the resolution between reference and noise-mitigated energies can be increased.

B. Six-excitation trial state

Due to the significant depth increase to the trial circuit, the six-excitation trial-state energies are expected to be more susceptible to noise than the four-excitation energies. On *ibmq_kolkata*, the noisier device, both $\langle \mathcal{H} \rangle$ and E_L fail to outperform the reference state. This is unexpected given that the reference-state calibration method should ensure that a trial state that outperforms the reference state in the absence of noise should remain competitive with the reference in the presence of noise. On the other hand, *ibm_torino* gives a value of $\langle \mathcal{H} \rangle$ that is significantly lower than the state-vector simulation, which is also unexpected. A possible explanation for both of these observations is that the noise-model assumptions used in error mitigation (Appendix C 4) may be invalid, possibly because of coherent errors in the state preparation. If this is the case, a solution would be to use randomized compiling [64] to convert coherent noise into unbiased stochastic noise. Alternatively, the unexpected behavior may be due to a failure to satisfy the N-representability conditions that determine whether the computed reduced density matrices are physical (see Appendix C 3).

III. DISCUSSION

While larger electronic structure computations have been performed, these are mostly made possible by encoding multiple spin orbitals into a single qubit [18,32–34] and/or through exploiting the assumption of weak entanglement [21,28]. This can allow simplification of paired double excitations to circuit elements involving only two CNOT or CZ gates, however, it prohibits the implementation of single and unpaired double excitations, limiting the effectiveness of the trial state. Here, knowledge of the initial state of the circuit is used to simplify the implementation of the first four excitations, however, the circuit is not restricted to paired coupled cluster states. After these four simplified excitations, the general form of the double-excitation gate is required, which results in the comparatively deep six-excitation circuit.

The accuracy obtained here for the water molecule compares well with the accuracy achieved for the benzene molecule using VQE for eight qubits and advanced circuit optimization on a trapped-ion quantum computer [35], though since the computations are performed for different molecules, a direct comparison is difficult. Additionally, the trapped-ion energy errors are reported relative to eight

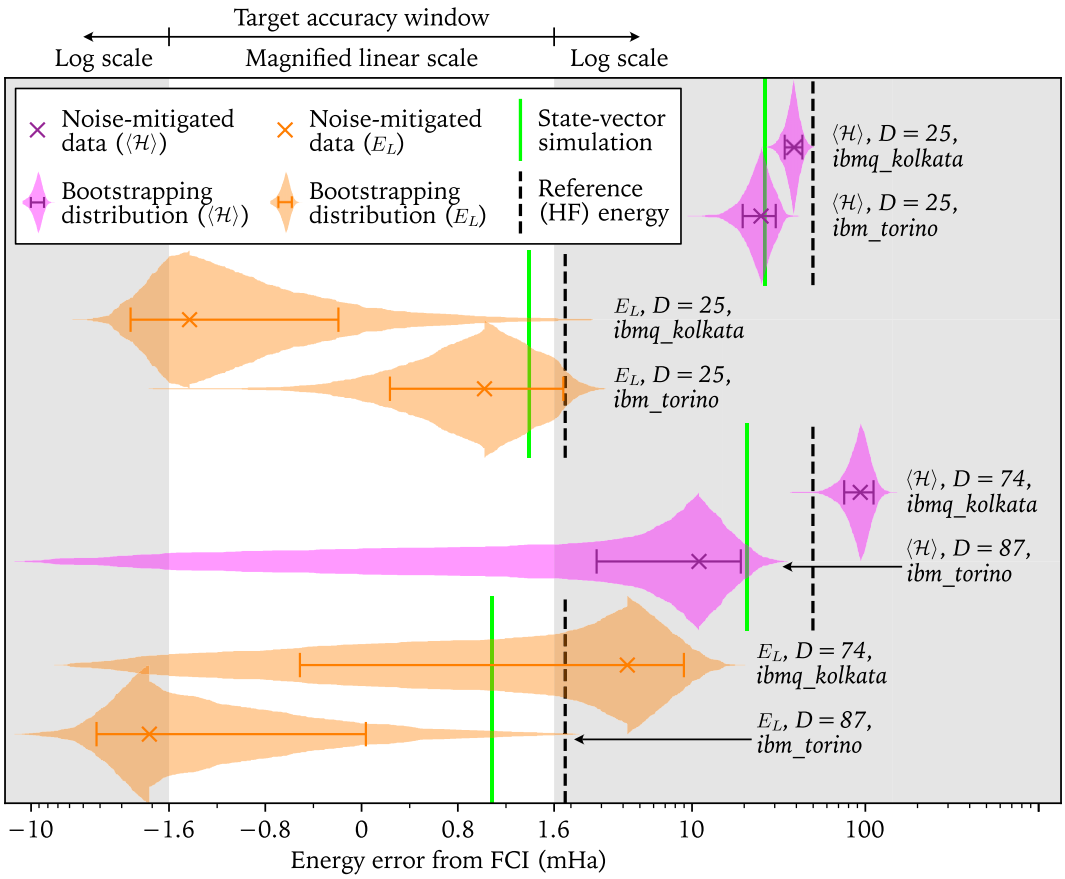

FIG. 3. Energy-estimation results for the water molecule based on measuring $\langle \mathcal{H} \rangle$ and E_L for the four-excitation ($D = 25$) and six-excitation ($D = 74$ and $D = 87$) circuits presented in Figs. 2(b)–2(e). The noiseless simulation results are shown for the trial state (in green) and the reference state (in dashed black). Noise-mitigated results, taken from the quantum devices *ibmq_kolkata* (10^5 shots per basis) and *ibm_torino* (2.5×10^4 shots per basis), are marked by magenta and orange crosses, respectively. The cumulative distributions indicate 500 statistical bootstrapping results and the error bars represent one standard deviation. Note that the horizontal axis uses a linear scale in the target accuracy window $[-1.6, 1.6]$ mHa and a logarithmic scale outside this window.

spin-orbital complete active space calculations, while we compare to the more accurate 14 spin-orbital FCI. Notably, the trapped-ion experiment required fewer state preparations at the expense of deeper circuits. This trade-off is favorable for trapped-ion qubits, which have longer coherence times but a lower repetition rate than superconducting qubits. The results obtained here also compare well with previous results for the F_2 molecule [36] computed using 12 superconducting qubits. While our results use fewer qubits, we achieve a higher accuracy. Additionally, the system considered here is at half-filling of the spin orbitals (where the problem is hardest) while moving away from half-filling (as in the case of F_2) makes the problem somewhat easier to solve.

We demonstrate that up to circuit depths >25 and widths of eight qubits, there is sufficient information contained in the output probability distributions from IBM Quantum hardware to outperform the classical Hartree-Fock reference state. It is useful to note that the energy scale on which this comparison takes place is much smaller than

TABLE I. Energy-estimation error in millihartree from FCI (−84.1813 Ha, excluding nuclear repulsion) for $\langle \mathcal{H} \rangle$ and E_L for different trial states and devices. The uncertainties are one standard deviation as calculated using 500 bootstrapping samples.

	Reference		Four excitation		Six excitation	
	$\langle \mathcal{H} \rangle$	E_L	$\langle \mathcal{H} \rangle$	E_L		
ibmq_kolkata			38.6 ± 4.4	-1.4 ± 1.2	93.5 ± 18.0	4.2 ± 4.7
ibm_torino			25.0 ± 5.4	1.0 ± 0.8	11.0 ± 8.2	-2.1 ± 2.1
State-vector simulation	49.8	1.8	26.2	1.4	20.8	1.1

the scale set by the maximally mixed state, which has error on the order of 1–2 Ha (500–800 mHa) for $\langle \mathcal{H} \rangle$ (E_L). This is 2–3 orders of magnitude larger than the energy scales of interest. This result demonstrates that it is possible for moments-based methods to extend the reach of variational algorithms for quantum chemistry on present-day and near-term quantum hardware, pointing the way towards a potential for quantum utility.

IV. METHODS

A. Electronic Hamiltonian

The electronic Hamiltonian is calculated using the *PySCF* [65] python package in the minimal STO-3G basis comprising 14 spin orbitals. The core orbitals (predominantly the 1*s* orbitals of the O atom) are frozen and the Hamiltonian moments are computed via Wick's theorem in the 12 spin-orbital basis. The Hamiltonian and its moments are then reduced further, to eight spin orbitals by freezing the four orbitals that participate least in the bonding. By computing moments in the 12 spin-orbital basis, the moments-corrected energies are able to extend the computation into the otherwise unused virtual orbitals. The FCI energies are also computed by *PySCF* in the 14 spin-orbital basis.

B. Trial circuit

The trial circuit is based on the Trotterization of the chemically inspired unitary coupled cluster doubles (UCCD) ansatz with the double excitations ordered based on their known contribution to the ground state. The circuit is then compiled to enforce linear connectivity between qubits and to respect the gateset native to the quantum device. See Appendix A for details.

C. Measurement bases

Using approximately 200 fermionic measurement bases allows computation of the 940 nontrivial elements of the four-body reduced density matrix. Each correlated excitation is decomposed to a product of uncorrelated excitations and these are grouped using a greedy algorithm. The measurement circuits are optimized by solving an integer linear program using *Gurobi* [66]. See Appendix B.

D. Energy minimization

Energy minimization is performed using a customized SPSA optimizer on noiseless, classical simulation. Minimization was performed 5 times using different seeds for the stochastic optimizer and the optimized parameters from the best run were used for the hardware demonstration, though the difference between runs was minimal. The optimization generally converged after 50–60 iterations.

E. IBM quantum backends

The original quantum computations in this work were performed on *ibmq_kolkata*, one of IBM's superconducting quantum devices [67], using 8 of the 27 available qubits. Later computations were performed on *ibm_torino* [44], a newer device with improved error rates and a different set of base gates. On *ibmq_kolkata* (*ibm_torino*), each measurement basis was measured 10^5 times (2.5×10^4 times); for 200 measurement bases at two different values of $\vec{\theta}$ (trial and reference states) and including the readout-noise calibration, this leads to approximately 42 million shots (10 million shots) and a total quantum computation time of 1.5 h (50 min) per trial state. Uncertainty estimation is performed by a statistical bootstrapping technique [68]. See Appendix D.

F. Noise mitigation

Scalable readout-noise mitigation, symmetry verification by postselection on the electron number and total spin, reduced density-matrix rescaling and reference-state calibration are implemented to mitigate noise accumulated during the quantum computation (see Appendix C). In addition the QCM method is able to partially account for the limited expressivity of the trial state and is robust to the remaining noise [22,41,42].

The data that support the findings of this study are available from the corresponding author upon reasonable request.

ACKNOWLEDGMENTS

The research was supported by the University of Melbourne through the establishment of the IBM Q Network Hub at the University. M.A.J. and H.J.V. are supported by the Australian Commonwealth Government through Research Training Program Scholarships. The authors would like to thank G. Mooney, H. Quiney, and A. Martin for useful discussions. This work was supported by resources provided by the Pawsey Supercomputing Research Centre with funding from the Australian Government and the Government of Western Australia, and by The University of Melbourne's Research Computing Services and the Petascale Campus Initiative.

L.C.L.H. conceived the project. M.A.J. set up the computational framework and performed the calculations and data analysis, with input from all authors.

The QCM method is the subject of an international patent application no. PCT/AU2021/050674.

APPENDIX A: TRIAL-STATE DESIGN

One of the most challenging aspects of implementing variational hybrid algorithms is choosing a trial circuit that is expressive enough to contain (a good approximation to) the true ground state and at the same time

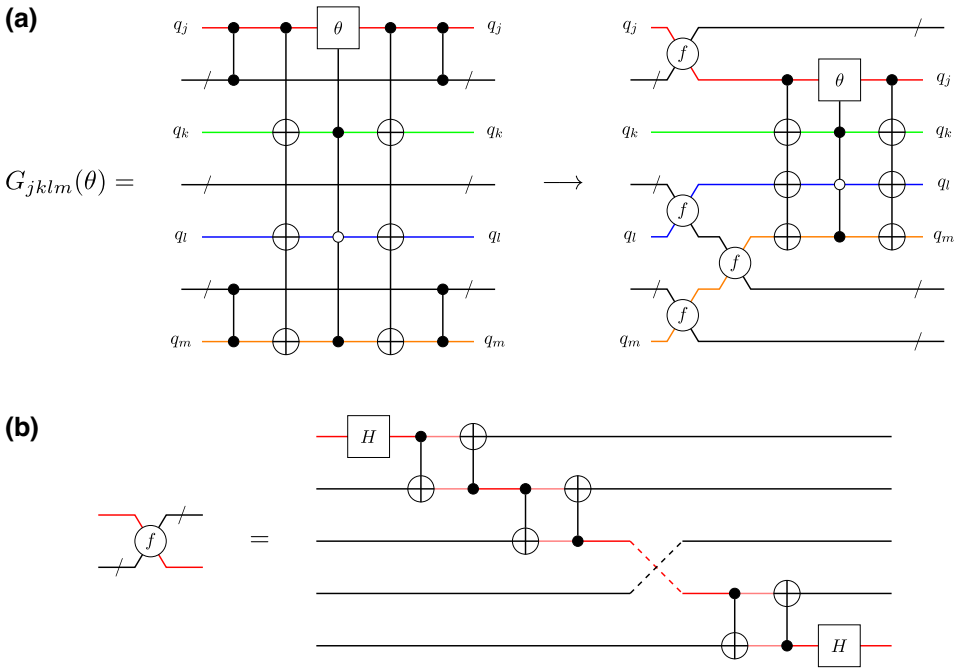

FIG. 4. (a) Conversion of the nonlocal fermionic double excitation to a local operation using fermionic SWAP operations, f, on linear connectivity. The θ gate is a Y rotation by angle θ and when combined with the CNOT gates, the controlled version can be decomposed to respect the linear connectivity as in Fig. 5. The CZ interactions are between each qubit in the bundle and one of the target qubits. The output of the two circuits is identical up to reordering with fermionic SWAPS. (b) Decomposition of the fermionic SWAP gate. Note that when many SWAPS are chained together in this way, the fermionic SWAP is cheaper to implement than the standard SWAP in both number of CNOTS and depth.

is shallow enough for the quantum device to remain coherent throughout the state preparation. In addition, care should be taken to avoid overparameterization of the ansatz, which can lead to so-called "barren plateaus" [69]—regions of exponentially vanishing gradient that cause classical optimization to become infeasible. A common, physically motivated, starting point for creating trial circuits for quantum chemistry is the unitary coupled cluster (UCC [43]) ansatz:

$$|\Psi_{\mathrm{UCC}}\rangle = e^{i\tau}|\Psi_{\mathrm{init}}\rangle, \qquad \tau = i(T - T^{\dagger}) \tag{A1}$$

where T is a weighted sum over all possible electronic excitations, including both uncorrelated (single) and correlated (double, triple, etc.) excitations. The initial state, $|\Psi_{\mathrm{init}}\rangle$, is an easily prepared state in the correct symmetry sector of Hilbert space and is usually taken to be the Hartree-Fock state, $|\Psi_{\mathrm{HF}}\rangle$. While the UCC trial state is expected to give a good approximation of the true ground state as the number of excitations is increased (in the limit that all excitations are included the trial state can exactly express the ground state), practical implementation requires both limitation to some maximum excitation order and Trotterization. Here, we use the double-excitation cluster operator, T_2, and Trotterize to first order:

$$|\Psi_{\mathrm{trial}}\rangle = \prod_{jklm} G_{jklm}(\theta_{jklm}) \times |\Psi_{\mathrm{HF}}\rangle, \qquad G_{jklm}(\theta) = \exp\left(\theta(a_j^{\dagger}a_k^{\dagger}a_l a_m - a_m^{\dagger}a_l^{\dagger}a_k a_j)\right), \tag{A2}$$

where $a_j^{\dagger}a_k^{\dagger}a_l a_m$ implements the electronic double excitation (the hermitian conjugate ensures unitarity of the G operator) and θ_{jklm} are variational parameters (corresponding to the weights in the original cluster operator, T). Given sufficient circuit depth, the single- and double-excitation ansatz has been seen to be an effective trial state [70]. In this case the exclusion of single-excitation operators is partially justified by the observation that leading single excitations cannot introduce multireference character and therefore cannot improve on the Hartree-Fock state, while trailing single excitations can be accounted for efficiently with classical processing [34,71,72]. In the Trotterization of the cluster operator, there is a degree of freedom as to the order in which noncommuting component excitations should be performed. Here, the order is chosen based on the known contribution of each excitation term to the true ground state [15] and, to limit circuit depth and permit implementation on present-day devices, only the first

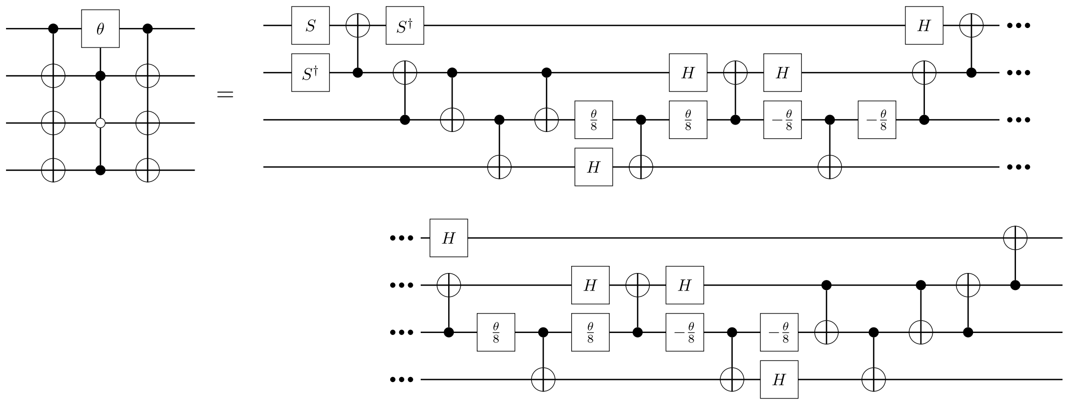

FIG. 5. Decomposition of the local double excitation to linear connectivity. The $\pm\theta/8$ gates are Y rotations by $\pm\theta/8$. This can be improved by prior knowledge of the input state, e.g., correlations between input qubits can be used to remove the long-range controls allowing for more efficient decomposition. Including the cost of performing fermionic SWAPS, the worst-case CNOT count is $4n+3$ for n qubits and spin orbitals. If it is required that the logical qubits return to their original physical qubit then the cost is $8n-13$ CNOT gates.

four–six excitations are considered. In practice this ordering would not be known *a priori*, however, it should be possible to make use of physical and chemical intuition and/or an adaptive ansatz [73,74] to determine a suitable ordering. This choice of ordering further justifies the lack of single excitations, as they are seen to have a smaller contribution than double excitations to the true ground state for the specific Hamiltonian considered here.

Since the trial-state preparation operator is now written as a product of double-excitations, a quantum circuit to implement the operator can be separated into a number of "blocks" each of which implements a double-excitation operation on a set of four fermionic modes. By considering the fermionic modes as logical qubits [75] on a linear array of physical qubits it can be shown that using the fermionic SWAP operator to permute the involved fermionic modes together converts the highly nonlocal excitation to a local interaction (shown in Fig. 4). Decomposition of the local fermionic interaction leads to 19 CNOT gates and 19 single-qubit [H, S, $S^\dagger$, and $R_y(\theta)$] gates with linear connectivity as shown in Fig. 5.

APPENDIX B: REDUCED DENSITY MATRICES

Generally, measurement within the VQE framework involves forming mutually commuting partitions of Hamiltonian fragments [7]. An alternative is to obtain a classical description of the quantum state from which relevant operators can be estimated, such as via "classical shadows" [76,77] or similar techniques. Here, the four-body reduced density matrix (4RDM) of the trial state is measured:

$$R_{jklm}^{j'k'l'm'} = \langle a_j^\dagger a_k^\dagger a_l^\dagger a_m^\dagger a_{j'} a_{k'} a_{l'} a_{m'} \rangle. \tag{B1}$$

The RDM is an efficient classical description of the state from which expectation values of lower-order excitations can be reconstructed:

$$R_{j_1..j_p}^{k_1...k_p} = \frac{1}{N_e - p} \sum_l R_{j_1..j_p l}^{k_1...k_p l}. \tag{B2}$$

Therefore, since any second quantized electronic structure Hamiltonian can be written in the form of Eq. (2), measurement of the 2RDM allows computation of *any* molecular energy (for a given trial circuit). Additionally, other properties of interest can be written as sums of low-order excitation operators and hence evaluated from the RDM. For example, the dipole moment is a sum over single-excitation operators

$$\mu = \sum_{jk} t_{jk} \langle a_j^\dagger a_k \rangle = \sum_{jk} t_{jk} R_j^k. \tag{B3}$$

In general, the fourth Hamiltonian moment contains up to eight-body excitations and therefore would require estimation of the 8RDM. However, due to the fact that there are only four electrons present in the trial state, excitation operators beyond fourth order will have vanishing expectation values.

Below we present a method for grouping and measuring correlated excitation operators for the estimation of reduced density matrices.

1. Filtering RDM elements

In general, the number of p RDM elements for N_s spin orbitals is

$$N_r = \frac{1}{2}\left(\frac{N_s!}{(N_s-p)!p!}\right)^2 = \mathcal{O}({N_s}^{2p}), \tag{B4}$$

where it can be shown that

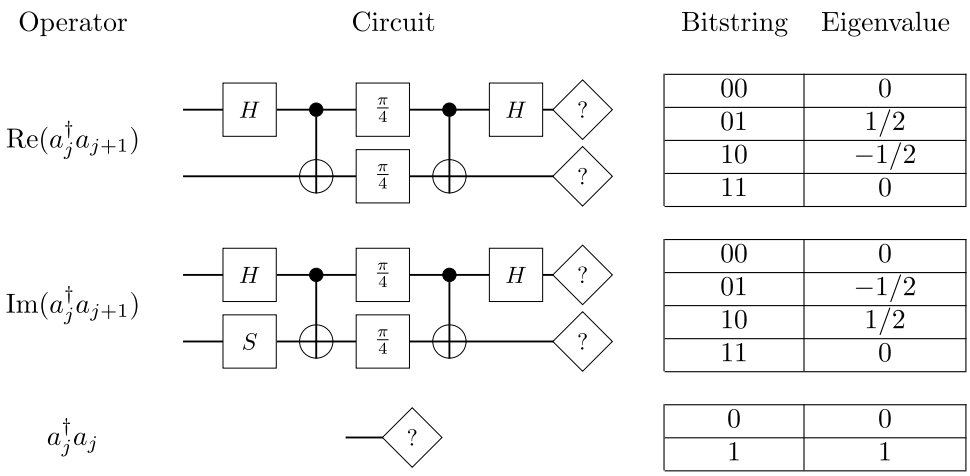

FIG. 6. Measurement circuits for single-excitation operators, $a_j^\dagger a_k$. When $k \notin \{j, j \pm 1\}$, fermionic SWAP operations can be used to reorder the logical qubits, then the circuits above can be applied. The $\pi/4$ gates are rotations about the y axis of the Bloch sphere by $\pi/4$.

(a) repeated indices in either the superscript or subscript on the left-hand side of Eq. (B1) lead to the right-hand side vanishing,

(b) permutations of indices introduce only factors of -1 due to fermionic anticommutation relations,

(c) interchanging upper and lower indices is equivalent to complex conjugation.

For the 4RDM considered here, this is 2485 operators but can be reduced further by considering spin symmetry [20]. Since the trial state is constructed within a specific spin subspace of the full Hilbert space, any excitation that does not conserve spin must have a vanishing expectation value. This restriction reduces the number of operators requiring measurement to 940, which (combined with the grouping method introduced below) is enough reduction to allow computation within a reasonable time.

2. Measurement of uncorrelated excitations

Before generalizing to correlated excitations, we first seek to measure uncorrelated excitations, i.e., operators of the form $a_j^\dagger a_k$. One method to do this involves applying the Jordan-Wigner transformation to convert the operator to a sum of Pauli strings and measuring each string separately. Unfortunately, this can lead to highly nonlocal strings, which are less likely to qubit-wise commute and are more prone to bit-flip errors in the readout process. Additionally, it is not possible to perform postselection when measuring Pauli strings as they do not necessarily commute with the total number operator or the spin-projection operator. Instead, consider the case $k = j \pm 1$. In this case, the excitation operator is a two-local operator and its real and imaginary components are diagonalized (and hence measured) by the gate sequences given in Fig. 6. For uncorrelated excitations on real-valued spin orbitals, the imaginary component is guaranteed to have vanishing expectation but will be required when evaluating correlated excitations. In the case that $k \notin \{j, j \pm 1\}$, the fermionic SWAP operation (Fig. 4) can be used to reorder the logical qubits so that they are adjacent. Finally, if $j = k$ then the operator is the number operator and its eigenvalue is simply the bit value when measured in the computational basis.

A disadvantage of this measurement scheme is the necessity of fermionic SWAP operations, which extends the circuit depth by $\mathcal{O}(N_s)$. On the other hand, it avoids nonlocal measurement strings and allows symmetry verification by postselection on the total electron number (since all circuits in Fig. 6 commute with the total electron number operator). Additionally, the only single excitations that do not commute with the total spin-projection operator are those for which j and k have different spins. From the arguments made in Appendix B 1, these operators vanish in expectation and their measurement is therefore unnecessary, allowing postselection based on the total spin projection.

3. Measurement of correlated excitations

Measuring correlated excitations directly is more difficult than measuring uncorrelated ones as the circuit depth required for diagonalization increases with excitation order. Instead, it is possible to decompose a correlated excitation into a sum of products of real and imaginary components of uncorrelated excitations, e.g., assuming unique subscripts:

$$
\begin{aligned}
a_{j_1}^\dagger a_{j_2}^\dagger a_{k_1} a_{k_2} &= -(a_{j_1}^\dagger a_{k_1})(a_{j_2}^\dagger a_{k_2}) \\
&= -\left(\mathrm{Re}(a_{j_1}^\dagger a_{k_1}) + i\,\mathrm{Im}(a_{j_1}^\dagger a_{k_1})\right)\left(\mathrm{Re}(a_{j_2}^\dagger a_{k_2}) + i\,\mathrm{Im}(a_{j_2}^\dagger a_{k_2})\right) \\
&= -\mathrm{Re}(a_{j_1}^\dagger a_{k_1})\mathrm{Re}(a_{j_2}^\dagger a_{k_2}) - i\mathrm{Re}(a_{j_1}^\dagger a_{k_1})\mathrm{Im}(a_{j_2}^\dagger a_{k_2}) \\
&\quad - i\mathrm{Im}(a_{j_1}^\dagger a_{k_1})\mathrm{Re}(a_{j_2}^\dagger a_{k_2}) + \mathrm{Im}(a_{j_1}^\dagger a_{k_1})\mathrm{Im}(a_{j_2}^\dagger a_{k_2}), \qquad \text{(B5)} \\
\langle a_{j_1}^\dagger a_{j_2}^\dagger a_{k_1} a_{k_2}\rangle &= -\left\langle \mathrm{Re}(a_{j_1}^\dagger a_{k_1})\mathrm{Re}(a_{j_2}^\dagger a_{k_2})\right\rangle + \left\langle \mathrm{Im}(a_{j_1}^\dagger a_{k_1})\mathrm{Im}(a_{j_2}^\dagger a_{k_2})\right\rangle, \qquad \text{(B6)}
\end{aligned}
$$

where the final step discards imaginary components that are expected to vanish in expectation for real-valued spin orbitals. Measurement of each term in the last line requires measuring the real and imaginary components (of different operators) simultaneously and multiplying their eigenvalues before averaging, in the same way that Pauli-string expectation values can be computed from individual Pauli measurements. The number of components in the sum grows as 2^p for an order p operator. However, since the maximum order is fixed, the decomposition contributes only a constant factor to the required number of measurements.

Since this method utilizes the same circuits as the uncorrelated method it can be made to allow postselection based on both the total number and total spin-projection operators. To ensure the latter is possible it is necessary to note that there are multiple ways to pair the indices in the first step of the decomposition. By pairing indices with the same spin (if this is not possible then the excitation must have vanishing expectation value by the arguments in Appendix B 1), the total spin-projection symmetry is preserved and postselection is possible. The decomposition can be generalized to cases where there are nonunique indices by first factoring out all nonunique indices and measuring these logical qubits in the computational basis.

4. Grouping of mutually commuting excitations

Since there are 940 nontrivial excitations to measure each of which requires up to eight different measurement bases, it is desirable to group mutually commuting excitations to be measured simultaneously. This grouping is performed by multiple applications of a greedy algorithm, a brief outline of which is given here, with an example given in Appendix E.

First we define a representation of a measurement basis as a list of qubit interactions. Each qubit can interact with one other qubit (or no other qubit for a number measurement $a_j^\dagger a_j$). For example, a measurement basis might be written as $(0,0),(1,2)$, indicating a number measurement on qubit 0 and a joint measurement on qubits 1 and 2. A basis measures an excitation if each creation index in the excitation is interacted with an annihilation index. Note that at this stage no distinction is made between the two different interactions given in Fig. 6.

The first stage of the algorithm partitions excitations into commuting (though not yet measurable) bases. The second stage partitions each basis further to allow measurement and can be mapped to the more well-known tensor-product-basis grouping problem.

Each partitioning is performed by iterating over the operators to be measured, selecting the "best" basis and updating it. If no compatible bases are found a new basis is appended to the set. Using this method the 940 excitations (with multiple components each) are partitioned into 200–204 measurement bases. Due to the greedy algorithm, the final number of bases is dependent on the order in which the excitations are presented to the algorithm and since the different trial states result in different arrangements of spin orbitals the ordering of excitations and hence number of bases is different between the two states.

5. Routing of logical qubits

An additional difficulty that has not been considered in the previous section is that of routing each pair of qubits together with as little circuit depth increase as possible. It is possible to show that each qubit can be interacted with each other qubit in linear depth [78], while the problem here is for each qubit to be interacted with a *specific* partner. Therefore the additional circuit depth should grow at most linearly with the number of qubits but naively routing qubit pairs sequentially often generates prohibitively deep circuit extensions. Additionally it is not hard to design an example for which the brickwork circuit [78] also performs poorly compared to solutions generated "by inspection." To solve this systematically we convert the problem to an integer linear program [79] and solve using Gurobi [66]. The binary variables we optimize are $x_{jk,t}^{(c)}$ and $y_{jk,t}^{(c)}$ where the x variables represent a logical qubit from pair c moving from physical qubit j to an adjacent (or the same) qubit k at time t and the y variables represent a logical qubit from pair c on physical qubit j interacting with its partner on adjacent qubit k at time t. Two constraints on the y variables immediately become apparent

(a) $y_{jk,t}^{(c)} = y_{kj,t}^{(c)}$, i.e., interactions must be between two logical qubits from the same pair, in the same timestep.

(b) $y_{jj,t}^{(c)} = 0$, i.e., two interacting logical qubits cannot be on the same physical qubit.

In addition we use the following constraints:

(1) Enforce starting positions of logical qubits ($S_j^{(c)}$ is 1 if a logical qubit from pair c starts on qubit j and 0 otherwise),

$$\sum_k \left[x_{jk,0}^{(c)} + y_{jk,0}^{(c)}\right] = S_j^{(c)}. \tag{B7}$$

(2) The number of logical qubits entering (or remaining on) a physical qubit at a given timestep must equal the number of logical qubits leaving (or remaining on) that physical qubit in the next timestep,

$$\sum_k \left[x_{jk,t}^{(c)} - x_{kj,t+1}^{(c)} + y_{jk,t}^{(c)} - y_{kj,t+1}^{(c)}\right] = 0. \tag{B8}$$

(3) The number of logical qubits entering a physical qubit at a given timestep must not be greater than 1,

$$\sum_{kc} \left[x_{jk,t}^{(c)} + y_{jk,t}^{(c)}\right] \leq 1. \tag{B9}$$

(4) If a logical qubit moves from physical qubit j to k at timestep t, then any logical qubit on k at t must move to j (i.e., enforce swapping behavior)

$$\sum_c \left[x_{jk,t}^{(c)} + \sum_{j' \neq j} x_{kj',t}^{(c)}\right] \leq 1. \tag{B10}$$

(5) Interactions must occur exactly once for each pair,

$$\sum_{jkt} \left[y_{jk,t}^{(c)}\right] = 2. \tag{B11}$$

Note that the conditions on y can be used to analytically simplify the implementation of the conditions above but have not been substituted here to maintain readability. Example outputs from the qubit-routing optimization are shown in Fig. 7.

APPENDIX C: NOISE MITIGATION

In the near term, fault-tolerant quantum computation (i.e., error correction) is not possible due to hardware constraints. Instead, various schemes have been proposed to mitigate hardware noise, usually at the cost of running additional circuits. The methods used in this work are detailed below.

1. Quantum readout-noise mitigation

One potential source of errors occurs during the qubit-readout process, where the qubit state may be misclassified. Quantum readout-error mitigation (QREM [80]) aims to address these readout errors by calibrating the readout procedure on known computational basis states. By creating and measuring a selection of computational basis states, the assignment matrix of each qubit can be formed. This matrix represents the transformation from the ideal output probability vector to the noisy vector. By inverting this matrix the transformation from noisy to ideal vector can be obtained. In the implementation used here [81], readout error on qubit j is averaged over (a subset of) the states of the other qubits but is not explicitly treated as being correlated noise. Therefore, the assignment matrices are 2×2 matrices and are efficiently invertible.

In practice QREM may lead to nonphysical quasiprobabilities. To enforce physicality the "negative probability" can be redistributed evenly between measurement outcomes with positive probability [82] (see also Ref. [83]).

2. Symmetry verification

Due to the physical motivation for the UCC-based trial state, the electron number and total electronic spin are fixed for any value of the trial-state parameters. Additionally, since we are able to measure both of these quantities simultaneously with the RDM elements through the fermionic measurement procedure described in Appendix B, they can be used for symmetry verification [84]. In other words, each time a bitstring is measured, if the corresponding electron number or total electronic spin is not the value expected from the trial-state design, then we can say with certainty that an error has occurred and discard the result. In this way only errors that commute with both symmetries are not detected and are able to influence the resulting energy.

Such a postselection strategy necessarily leads to a reduction in the effective number of counts. In the demonstrations performed here, only 56% and 22% of measurement shots performed on *ibmq_kolkata* are accepted for the four- and six-excitation trial circuits, respectively, while the numbers for *ibm_torino* are 63% and 46%. Note that generating bitstrings at random would have a 14% acceptance rate.

3. Reduced density-matrix rescaling

The reduced density matrices as given by Eq. (B1) are only well defined for states with a fixed number of electrons, otherwise the use of Eq. (B2) can lead to rapid accumulation of error. Unfortunately, and somewhat counterintuitively, symmetry verification alone is not sufficient to enforce the correct electron number for the RDM. This is because symmetry verification enforces the correct correlations only between elements of the same measurement basis, so properties that are reconstructed from measurements in multiple bases can violate symmetry constraints. To resolve this, it can be noted that the electron number

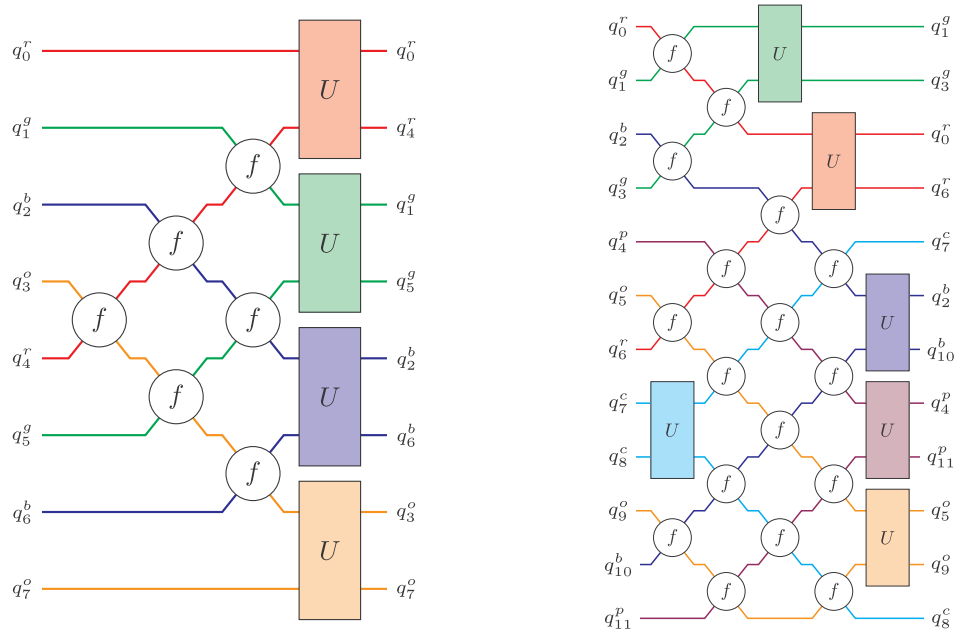

FIG. 7. Example circuits for logical-qubit routing from the integer linear program for 8 (left) and 12 (right) qubits. The problem is to interact each qubit with its partner (matching color and superscript) in as few timesteps as possible, with a secondary objective to minimize the number of SWAPS. The f gates are fermionic SWAPS while the U gates are the interactions given in Fig. 6 for converting the joint measurement into single-qubit measurements.

can be directly related to the trace of the ideal p RDM by

$$\begin{aligned}\mathrm{Tr}({}_{\text{ideal}}R) &\equiv \sum_{j_1<j_2\cdots<j_p} {}_{\text{ideal}}R^{j_1 j_2 \cdots j_p}_{j_1 j_2 \cdots j_p} \\ &= \frac{N_e!}{p!(N_e-p)!},\end{aligned} \tag{C1}$$

and to enforce the correct trace (and therefore electron number) it is possible to simply rescale the RDM by the ratio of the ideal trace to the measured trace [20], i.e.,

$${}_{\text{corrected}}R = \frac{N_e!}{p!(N_e-p)!}\frac{1}{\mathrm{Tr}({}_{\text{noisy}}R)} \times {}_{\text{noisy}}R. \tag{C2}$$

We note that the correct trace is a necessary, but not sufficient, condition for the p RDM to represent a valid physical state. In general, the RDM needs to satisfy the N-representability conditions [85–89].

Using $\mathbf{j}$ to represent a list of indices, $j_1 j_2 \ldots$, for brevity, then a (still insufficient) list of necessary conditions for validity of the p RDM is [88]

(1) Hermiticity:

$$R^{\mathbf{k}}_{\mathbf{j}} = (R^{\mathbf{j}}_{\mathbf{k}})^*, \tag{C3}$$

(2) Antisymmetry:

$$\begin{aligned}R^{\mathbf{k}}_{P(\mathbf{j})} &= \sigma_P R^{\mathbf{j}}_{\mathbf{k}}, \\ R^{P(\mathbf{k})}_{\mathbf{j}} &= \sigma_P R^{\mathbf{j}}_{\mathbf{k}},\end{aligned} \tag{C4}$$

where $P(\mathbf{j})$ is a permutation of the indices contained in $\mathbf{j}$ and $\sigma_P = \pm 1$ is the sign of the permutation.

(3) Contraction:

$$R^{k_1 \ldots k_p}_{j_1 \ldots j_p} = \frac{1}{N_e - p}\sum_l R^{k_1 \ldots k_p l}_{j_1 \ldots j_p l}, \tag{C5}$$

(4) Trace:

$$\mathrm{Tr}(R) = \frac{N_e!}{p!(N_e-p)!}, \tag{C6}$$

(5) Positive semidefiniteness:

$$R \succeq 0. \tag{C7}$$

Further constraints can be derived from (for example) the requirement that the trial state is an eigenstate of the spin-projection operator [86,88]. Note that in this work, the hermiticity and antisymmetry conditions are used to reduce the number of required measurements in Appendix B 1, the contraction condition is exactly Eq. (B2), which is used to extract lower-order expectation values from high-order

RDMs, and the trace condition is enforced by the rescaling technique. In future work, additional constraints could potentially be implemented to further detect and mitigate errors as has been proposed [88] and performed [14] for 2RDMs.

An interesting feature of measuring RDMs is that even with this rescaling technique, if no other error mitigation is applied, the energy estimates it produces have lower accuracy than estimating the energy directly from counts. On the other hand, we see that the variance of the energy estimates is significantly reduced by the RDM method. This is advantageous because many error-mitigation techniques improve accuracy at the cost of an increased variance. In combination (in particular, with the noise calibration described below), the RDM method results in both more accurate and more precise (lower-variance) estimates than calculating energy estimates directly from counts.

4. Reference-state noise calibration

Another powerful error-mitigation technique is based on the existence of reference circuits [27,90], circuits with similar overall structure to the trial state but with a limited number of non-Clifford gates. The simplest reference state for the circuit used here is the Hartree-Fock state, which can be prepared by setting all the variational parameters to 0. Since properties of the Hartree-Fock state can be efficiently evaluated classically, comparison between classical and quantum computations can be used to "learn" the noise model of the device. Here, a global white-noise model is assumed,

$$\begin{aligned} \rho \to \rho^{\text{noisy}} &= (1-q)\rho + qJ, \\ \Rightarrow \text{Tr}(\mathcal{H}\rho^{\text{noisy}}) &= (1-q)\text{Tr}(\mathcal{H}\rho) + q\text{Tr}(\mathcal{H}J), \end{aligned} \tag{C8}$$

where q represents the effective global error rate and J is related to the maximally mixed state (concretely it is the image of the maximally mixed state after application of other error-mitigation techniques, e.g., postselection will first zero out some entries of the mixed state, $I/2^{N_s}$, then renormalize).

This model has been seen to be reasonable for certain classes of circuits [91] and circuits could be made to better obey the model through randomized compiling techniques [64]. Using the classical (noiseless) and quantum (noisy) expectation values from the Hartree-Fock state, an estimate, $\hat{q}$, for the error rate can be obtained by solving Eq. (C8). For the two-excitation trial-state circuits, the calculated noise level on *ibmq_kolkata* (*ibm_torino*) ranges between 3–40% (5–20%), depending on the application of other noise-mitigation techniques. Following this, quantum computation of properties of the trial state can be corrected using the noise estimate, $\hat{q}$. A graphical representation of this process is shown in Fig. 8.

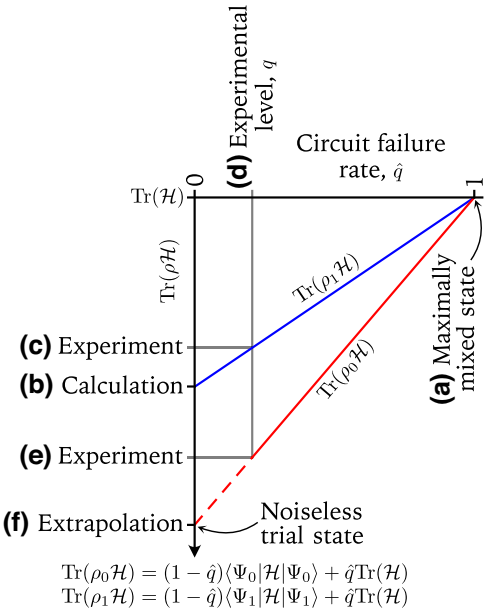

FIG. 8. The reference-state calibration method can be visualized on a plot of expectation value versus circuit failure rate. The maximally noisy (a) and noiseless (b) energies of the reference state are classically tractable, so using quantum computation of the reference energy (c) allows estimation of the effective noise level in the device (d). By evaluating the energy of the trial state on a quantum computer (e) the noiseless energy of the trial state (f) can be estimated by extrapolation. The linear relationship is due to the white-noise model, with more sophisticated noise models requiring more reference-state calculations to fit [90]. The reference-state error-mitigation scheme differs from the usual zero-noise extrapolation method [11,92] since the additional information is obtained through reference states instead of artificially increasing the noise.

APPENDIX D: STATISTICAL BOOTSTRAPPING

Standard deviations are estimated using a bootstrapping method [68]. By resampling (classically) 500×10^5 ($500 \times 2.5 \times 10^4$ for *ibm_torino*) times from the probability distributions obtained from the quantum hardware, a distribution of energy estimates is obtained from which error estimates can be extracted.

APPENDIX E: BASIS GROUPING EXAMPLE

1. Fermionic measurement grouping (first level)

Consider measuring the spin-conserving excitations for the 2RDM on four spin orbitals:

$$\begin{array}{llll} a_0^\dagger a_1^\dagger a_0 a_1 & & & \\ & a_0^\dagger a_2^\dagger a_0 a_2 & a_0^\dagger a_2^\dagger a_1 a_2 & a_0^\dagger a_2^\dagger a_0 a_3 \quad a_0^\dagger a_2^\dagger a_1 a_3 \\ & & a_1^\dagger a_2^\dagger a_1 a_2 & a_1^\dagger a_2^\dagger a_0 a_3 \quad a_1^\dagger a_2^\dagger a_1 a_3 \\ & & & a_0^\dagger a_3^\dagger a_0 a_3 \quad a_0^\dagger a_3^\dagger a_1 a_3 \\ & & & a_1^\dagger a_3^\dagger a_1 a_3 \\ & & & a_2^\dagger a_3^\dagger a_2 a_3, \end{array}$$

where spin orbitals 0 and 1 have spin up and spin orbitals 2 and 3 have spin down. We begin with an empty basis set, {}, and for each excitation above, search for a commuting basis. If a basis is found we update it, if no basis is found we add a new one.

(1) $a_0^\dagger a_1^\dagger a_0 a_1$
   (a) $\{\} \rightarrow \{(0,1)\}$.
   (b) No commuting basis. New basis $(0,1)$.

(2) $a_0^\dagger a_2^\dagger a_0 a_2$
   (a) $\{(0,1)\} \rightarrow \begin{Bmatrix} (0,1) \\ (0,0),(2,2) \end{Bmatrix}$.
   (b) No commuting basis. New basis $(0,0),(2,2)$. Note that spin orbitals 0 and 2 have opposite spin, so interacting them is not allowed (fails to commute with the spin-projection operator).

(3) $a_0^\dagger a_2^\dagger a_1 a_2$
   (a) $\begin{Bmatrix} (0,1) \\ (0,0),(2,2) \end{Bmatrix} \rightarrow \begin{Bmatrix} (0,1),(2,2) \\ (0,0),(2,2) \end{Bmatrix}$.
   (b) Commuting basis, $(0,1)$ found and updated to $(0,1),(2,2)$.

(4) $a_0^\dagger a_2^\dagger a_0 a_3$
   (a) $\begin{Bmatrix} (0,1),(2,2) \\ (0,0),(2,2) \end{Bmatrix} \rightarrow \begin{Bmatrix} (0,1),(2,2) \\ (0,0),(2,2) \\ (0,0),(2,3) \end{Bmatrix}$.
   (b) No commuting basis. New basis $(0,0),(2,3)$.

(5) $a_0^\dagger a_2^\dagger a_1 a_3$
   (a) $\begin{Bmatrix} (0,1),(2,2) \\ (0,0),(2,2) \\ (0,0),(2,3) \end{Bmatrix} \rightarrow \begin{Bmatrix} (0,1),(2,2) \\ (0,0),(2,2) \\ (0,0),(2,3) \\ (0,1),(2,3) \end{Bmatrix}$.
   (b) No commuting basis. New basis $(0,1),(2,3)$.

(6) $a_1^\dagger a_2^\dagger a_1 a_2$
   (a) $\begin{Bmatrix} (0,1),(2,2) \\ (0,0),(2,2) \\ (0,0),(2,3) \\ (0,1),(2,3) \end{Bmatrix} \rightarrow \begin{Bmatrix} (0,1),(2,2) \\ (0,0),(1,1),(2,2) \\ (0,0),(2,3) \\ (0,1),(2,3) \end{Bmatrix}$.
   (b) Commuting basis, $(0,0),(2,2)$ found and updated to $(0,0),(1,1),(2,2)$.

(7) $a_1^\dagger a_2^\dagger a_0 a_3$
   (a) $\begin{Bmatrix} (0,1),(2,2) \\ (0,0),(1,1),(2,2) \\ (0,0),(2,3) \\ (0,1),(2,3) \end{Bmatrix}$.
   (b) Measuring basis $(0,1),(2,3)$ found.

(8) $a_1^\dagger a_2^\dagger a_1 a_3$
   (a) $\begin{Bmatrix} (0,1),(2,2) \\ (0,0),(1,1),(2,2) \\ (0,0),(2,3) \\ (0,1),(2,3) \end{Bmatrix} \rightarrow \begin{Bmatrix} (0,1),(2,2) \\ (0,0),(1,1),(2,2) \\ (0,0),(1,1),(2,3) \\ (0,1),(2,3) \end{Bmatrix}$.
   (b) Commuting basis $(0,0),(2,3)$ found and updated to $(0,0),(1,1),(2,3)$.

(9) $a_0^\dagger a_3^\dagger a_0 a_3$
   (a) $\begin{Bmatrix} (0,1),(2,2) \\ (0,0),(1,1),(2,2) \\ (0,0),(1,1),(2,3) \\ (0,1),(2,3) \end{Bmatrix} \rightarrow$.
   $\begin{Bmatrix} (0,1),(2,2) \\ (0,0),(1,1),(2,2),(3,3) \\ (0,0),(1,1),(2,3) \\ (0,1),(2,3) \end{Bmatrix}$.
   (b) Commuting basis $(0,0),(1,1),(2,2)$ found and updated to $(0,0),(1,1),(2,2),(3,3)$.

(10) $a_0^\dagger a_3^\dagger a_1 a_3$
   (a) $\begin{Bmatrix} (0,1),(2,2) \\ (0,0),(1,1),(2,2),(3,3) \\ (0,0),(1,1),(2,3) \\ (0,1),(2,3) \end{Bmatrix} \rightarrow$
   $\begin{Bmatrix} (0,1),(2,2),(3,3) \\ (0,0),(1,1),(2,2),(3,3) \\ (0,0),(1,1),(2,3) \\ (0,1),(2,3) \end{Bmatrix}$.
   (b) Commuting basis $(0,1),(2,2)$ found and updated to $(0,1),(2,2),(3,3)$

(11) $a_1^\dagger a_3^\dagger a_1 a_3$
   (a) $\begin{Bmatrix} (0,1),(2,2),(3,3) \\ (0,0),(1,1),(2,2),(3,3) \\ (0,0),(1,1),(2,3) \\ (0,1),(2,3) \end{Bmatrix}$.
   (b) Measuring basis $(0,0),(1,1),(2,2),(3,3)$ found.

(12) $a_2^\dagger a_3^\dagger a_2 a_3$
   (a) $\begin{Bmatrix} (0,1),(2,2),(3,3) \\ (0,0),(1,1),(2,2),(3,3) \\ (0,0),(1,1),(2,3) \\ (0,1),(2,3) \end{Bmatrix}$.
   (b) Measuring basis $(0,1),(2,2),(3,3)$ found. Note that $(0,0),(1,1),(2,2),(3,3)$ also measures this excitation. At this point the choice between the two is arbitrary.

2. Mapping to Pauli measurement grouping (second level)

Based on the results for the previous example, we now need to assign the interaction types (i.e., which circuit from Fig. 6) to use for each interaction in each basis in the set. The assignment of interactions in a given basis are independent of the other bases and so we can take the bases in any order.

(1) $(0,0),(1,1),(2,2),(3,3)$
   (a) Measures $a_1^\dagger a_3^\dagger a_1 a_3$, $a_0^\dagger a_3^\dagger a_0 a_3$, $a_1^\dagger a_2^\dagger a_1 a_2$ and $a_0^\dagger a_2^\dagger a_0 a_2$.
   (b) Since each qubit is "interacted" with itself, the only choice is to measure in the computational basis. This results in measuring the number operators $a_j^\dagger a_j$

(for $j \in [0,3]$) from which the products can be reconstructed. In this case no further partitioning is required.

(2) $(0,1),(2,2),(3,3)$

(a) Measures $a_2^\dagger a_3^\dagger a_2 a_3$, $a_0^\dagger a_3^\dagger a_1 a_3$, $a_0^\dagger a_2^\dagger a_1 a_2$ and $a_0^\dagger a_1^\dagger a_0 a_1$.

(b) Only one of the interactions is between different qubits, therefore this can be mapped to a (trivial) one-site tensor-product basis-grouping problem. We will use "site" for these virtual qubits and "qubits" for the actual qubits.

(i) The interaction $(0,1)$ is mapped to the first (and only) site. We need to decide what basis to measure it in.

(ii) The excitation operators are mapped to the Pauli operators to be grouped.

(A) $a_2^\dagger a_3^\dagger a_2 a_3$ does not act on the site, so maps to I.

(B) $a_0^\dagger a_3^\dagger a_1 a_3$ acts on the site as $a_0^\dagger a_1 \to \mathrm{Re}(a_0^\dagger a_1) \to X$.

(C) $a_0^\dagger a_2^\dagger a_1 a_2$ also maps to X.

(D) $a_0^\dagger a_1^\dagger a_0 a_1$ is a special case and can be extracted from either of the interactions in Fig. 6 (since they are both number conserving by design) so we can also map this to I.

(iii) As stated above, the grouping problem in this case is trivial, since there is only one site and only one nonidentity operator (X).

(3) $(0,0),(1,1),(2,3)$

(a) Measures $a_1^\dagger a_2^\dagger a_1 a_3$, $a_0^\dagger a_2^\dagger a_0 a_3$.

(b) Similar to above, this maps to the problem of one site and one operator (X).

(4) $(0,1),(2,3)$

(a) Measures $a_1^\dagger a_2^\dagger a_0 a_3$ and $a_0^\dagger a_2^\dagger a_1 a_3$.

(b) In this case we map to a two-site tensor-product basis problem

(i) $(0,1)$ and $(2,3)$ become the first and second sites, respectively.

(ii) The excitation operators map as

(A) $a_1^\dagger a_2^\dagger a_0 a_3 \to \mathrm{Re}(a_1^\dagger a_0)\mathrm{Re}(a_2^\dagger a_3) + \mathrm{Im}(a_1^\dagger a_0)\mathrm{Im}(a_2^\dagger a_3) \to XX + YY$.

(B) $a_0^\dagger a_2^\dagger a_1 a_3 \to \mathrm{Re}(a_0^\dagger a_1)\mathrm{Re}(a_2^\dagger a_3) + \mathrm{Im}(a_0^\dagger a_1)\mathrm{Im}(a_2^\dagger a_3) \to XX + YY$.

(C) This then leads to the problem of finding a tensor product grouping for XX and YY (which trivially requires two bases)

In nontrivial cases, i.e., for a higher number of spin orbitals, the Pauli-measurement grouping problem is solved using a greedy algorithm analogous to the Fermionic grouping algorithm of the previous section.

[1] A. Y. Kitaev, Quantum measurements and the Abelian stabilizer problem, ArXiv:quant-ph/9511026.

[2] M. Reiher, N. Wiebe, K. M. Svore, D. Wecker, and M. Troyer, Elucidating reaction mechanisms on quantum computers, Proc. Natl. Acad. Sci. **114**, 7555 (2017).

[3] J. Preskill, Quantum computing in the NISQ era and beyond, Quantum **2**, 79 (2018).

[4] A. Peruzzo, J. McClean, P. Shadbolt, M.-H. Yung, X.-Q. Zhou, P. J. Love, A. Aspuru-Guzik, and J. L. O'Brien, A variational eigenvalue solver on a photonic quantum processor, Nat. Commun. **5**, 4213 (2014).

[5] M. Cerezo, A. Arrasmith, R. Babbush, S. C. Benjamin, S. Endo, K. Fujii, J. R. McClean, K. Mitarai, X. Yuan, L. Cincio, and P. J. Coles, Variational quantum algorithms, Nat. Rev. Phys. **3**, 625 (2021).

[6] P. J. J. O'Malley, R. Babbush, I. D. Kivlichan, J. Romero, J. R. McClean, R. Barends, J. Kelly, P. Roushan, A. Tranter, N. Ding, *et al.*, Scalable quantum simulation of molecular energies, Phys. Rev. X **6**, 031007 (2016).

[7] A. Kandala, A. Mezzacapo, K. Temme, M. Takita, M. Brink, J. M. Chow, and J. M. Gambetta, Hardware-efficient variational quantum eigensolver for small molecules and quantum magnets, Nature **549**, 242 (2017).

[8] Y. Shen, X. Zhang, S. Zhang, J.-N. Zhang, M.-H. Yung, and K. Kim, Quantum implementation of the unitary coupled cluster for simulating molecular electronic structure, Phys. Rev. A **95**, 020501(R) (2017).

[9] J. I. Colless, V. V. Ramasesh, D. Dahlen, M. S. Blok, M. E. Kimchi-Schwartz, J. R. McClean, J. Carter, W. A. de Jong, and I. Siddiqi, Computation of molecular spectra on a quantum processor with an error-resilient algorithm, Phys. Rev. X **8**, 011021 (2018).

[10] M. Ganzhorn, D. J. Egger, P. Barkoutsos, P. Ollitrault, G. Salis, N. Moll, M. Roth, A. Fuhrer, P. Mueller, S. Woerner, I. Tavernelli, and S. Filipp, Gate-efficient simulation of molecular eigenstates on a quantum computer, Phys. Rev. Appl. **11**, 044092 (2019).

[11] A. Kandala, K. Temme, A. D. Córcoles, A. Mezzacapo, J. M. Chow, and J. M. Gambetta, Error mitigation extends the computational reach of a noisy quantum processor, Nature **567**, 491 (2019).

[12] A. J. McCaskey, Z. P. Parks, J. Jakowski, S. V. Moore, T. D. Morris, T. S. Humble, and R. C. Pooser, Quantum chemistry as a benchmark for near-term quantum computers, npj Quantum Inf. **5**, 99 (2019).

[13] R. Sagastizabal, X. Bonet-Monroig, M. Singh, M. A. Rol, C. C. Bultink, X. Fu, C. H. Price, V. P. Ostroukh, N. Muthusubramanian, A. Bruno, M. Beekman, N. Haider, T. E. O'Brien, and L. DiCarlo, Experimental error mitigation via symmetry verification in a variational quantum eigensolver, Phys. Rev. A **100**, 010302(R) (2019).

[14] S. E. Smart and D. A. Mazziotti, Quantum-classical hybrid algorithm using an error-mitigating n-representability condition to compute the Mott metal-insulator transition, Phys. Rev. A **100**, 022517 (2019).

[15] Y. Nam, J.-S. Chen, N. C. Pisenti, K. Wright, C. Delaney, D. Maslov, K. R. Brown, S. Allen, J. M. Amini, J. Apisdorf, *et al.*, Ground-state energy estimation of the water molecule on a trapped-ion quantum computer, npj Quantum Inf. **6**, 33 (2020).

[16] S. E. Smart and D. A. Mazziotti, Efficient two-electron ansatz for benchmarking quantum chemistry on a quantum computer, Phys. Rev. Res. **2**, 023048 (2020).
[17] Q. Gao, H. Nakamura, T. P. Gujarati, G. O. Jones, J. E. Rice, S. P. Wood, M. Pistoia, J. M. Garcia, and N. Yamamoto, Computational investigations of the lithium superoxide dimer rearrangement on noisy quantum devices, J. Phys. Chem. A **125**, 1827 (2021).
[18] Y. Kawashima, E. Lloyd, M. P. Coons, Y. Nam, S. Matsuura, A. J. Garza, S. Johri, L. Huntington, V. Senicourt, A. O. Maksymov, *et al.*, Optimizing electronic structure simulations on a trapped-ion quantum computer using problem decomposition, Nat. Commun. Phys. **4**, 245 (2021).
[19] J. E. Rice, T. P. Gujarati, M. Motta, T. Y. Takeshita, E. Lee, J. A. Latone, and J. M. Garcia, Quantum computation of dominant products in lithium-sulfur batteries, J. Chem. Phys. **154**, 134115 (2021).
[20] J. Tilly, P. V. Sriluckshmy, A. Patel, E. Fontana, I. Rungger, E. Grant, R. Anderson, J. Tennyson, and G. H. Booth, Reduced density matrix sampling: Self-consistent embedding and multiscale electronic structure on current generation quantum computers, Phys. Rev. Res. **3**, 033230 (2021).
[21] A. Eddins, M. Motta, T. P. Gujarati, S. Bravyi, A. Mezzacapo, C. Hadfield, and S. Sheldon, Doubling the size of quantum simulators by entanglement forging, PRX Quantum **3**, 010309 (2022).
[22] M. A. Jones, H. J. Vallury, C. D. Hill, and L. C. L. Hollenberg, Chemistry beyond the Hartree–Fock energy via quantum computed moments, Sci. Rep. **12**, 8985 (2022).
[23] Y. Shee, P.-K. Tsai, C.-L. Hong, H.-C. Cheng, and H.-S. Goan, Qubit-efficient encoding scheme for quantum simulations of electronic structure, Phys. Rev. Res. **4**, 023154 (2022).
[24] P. Suchsland, F. Tacchino, M. H. Fischer, T. Neupert, P. K. Barkoutsos, and I. Tavernelli, Algorithmic error mitigation scheme for current quantum processors, Quantum **5**, 492 (2021).
[25] K. Yamamoto, D. Z. Manrique, I. T. Khan, H. Sawada, and D. M. n. Ramo, Quantum hardware calculations of periodic systems with partition-measurement symmetry verification: Simplified models of hydrogen chain and iron crystals, Phys. Rev. Res. **4**, 033110 (2022).
[26] I. T. Khan, M. Tudorovskaya, J. J. M. Kirsopp, D. Muñoz Ramo, P. Warrier, D. K. Papanastasiou, and R. Singh, Chemically aware unitary coupled cluster with ab initio calculations on an ion trap quantum computer: A refrigerant chemicals' application, J. Chem. Phys. **158**, 214114 (2023).
[27] P. Lolur, M. Skogh, W. Dobrautz, C. Warren, J. Biznárová, A. Osman, G. Tancredi, G. Wendin, J. Bylander, and M. Rahm, Reference-state error mitigation: A strategy for high accuracy quantum computation of chemistry, J. Chem. Theory Comput. **19**, 783 (2023).
[28] M. Motta, G. O. Jones, J. E. Rice, T. P. Gujarati, R. Sakuma, I. Liepuoniute, J. M. Garcia, and Y.-Y. Ohnishi, Quantum chemistry simulation of ground- and excited-state properties of the sulfonium cation on a superconducting quantum processor, Chem. Sci. **14**, 2915 (2023).
[29] M. Rossmannek, F. Pavošević, A. Rubio, and I. Tavernelli, Quantum embedding method for the simulation of strongly correlated systems on quantum computers, J. Phys. Chem. Lett. **14**, 3491 (2023).
[30] S. Shirai, H. Iwakiri, K. Kanno, T. Horiba, K. Omiya, H. Hirai, and S. Koh, Computational analysis of chemical reactions using a variational quantum eigensolver algorithm without specifying spin multiplicity, ACS Omega **8**, 19917 (2023).
[31] D. Yoffe, A. Natan, and A. Makmal, A qubit-efficient variational selected configuration-interaction method, ArXiV: 2302.06691.
[32] Google AI Quantum, Hartree-Fock on a superconducting qubit quantum computer, Science **369**, 1084 (2020).
[33] T. E. O'Brien, G. Anselmetti, F. Gkritsis, V. E. Elfving, S. Polla, W. J. Huggins, O. Oumarou, K. Kechedzhi, D. Abanin, R. Acharya, *et al.*, Purification-based quantum error mitigation of pair-correlated electron simulations, Nat. Phys. **19**, 1787 (2023).
[34] L. Zhao, J. Goings, K. Shin, W. Kyoung, J. I. Fuks, J.-K. Kevin Rhee, Y. M. Rhee, K. Wright, J. Nguyen, J. Kim, and S. Johri, Orbital-optimized pair-correlated electron simulations on trapped-ion quantum computers, npj Quantum Inf. **9**, 60 (2023).
[35] J. Goings, L. Zhao, J. Jakowski, T. Morris, and R. Pooser, in *2023 IEEE International Conference on Quantum Computing and Engineering (QCE)* (IEEE Computer Society, Los Alamitos, CA, USA, 2023), p. 76.
[36] S. Guo, J. Sun, H. Qian, M. Gong, Y. Zhang, F. Chen, Y. Ye, Y. Wu, S. Cao, K. Liu, *et al.*, Experimental quantum computational chemistry with optimised unitary coupled cluster ansatz, ArXiv:2212.08006.
[37] L. C. L. Hollenberg, Plaquette expansion in lattice Hamiltonian models, Phys. Rev. D **47**, 1640 (1993).
[38] L. C. L. Hollenberg and N. S. Witte, General nonperturbative estimate of the energy density of lattice hamiltonians, Phys. Rev. D **50**, 3382 (1994).
[39] L. C. L. Hollenberg and N. S. Witte, Analytic solution for the ground-state energy of the extensive many-body problem, Phys. Rev. B **54**, 16309 (1996).
[40] L. C. L. Hollenberg, D. C. Bardos, and N. S. Witte, Lanczos cluster expansion for non-extensive systems, Z. Phys. D **38**, 249 (1996).
[41] H. J. Vallury, M. A. Jones, C. D. Hill, and L. C. L. Hollenberg, Quantum computed moments correction to variational estimates, Quantum **4**, 373 (2020).
[42] H. J. Vallury, M. A. Jones, G. A. L. White, F. M. Creevey, C. D. Hill, and L. C. L. Hollenberg, Noise-robust ground state energy estimates from deep quantum circuits, Quantum **7**, 1109 (2023).
[43] A. G. Taube and R. J. Bartlett, New perspectives on unitary coupled-cluster theory, Int. J. Quantum Chem. **106**, 3393 (2006).
[44] See Supplemental Material at http://link.aps.org/supplemental/10.1103/PhysRevApplied.21.064017 for detailed circuits and characteristics of the qubits used on *ibm_torino*.
[45] P. Jordan and E. Wigner, Über das paulische äquivalenzverbot, Z. Phys. **47**, 631 (1928).

[46] J. R. McClean, M. E. Kimchi-Schwartz, J. Carter, and W. A. de Jong, Hybrid quantum-classical hierarchy for mitigation of decoherence and determination of excited states, Phys. Rev. A **95**, 042308 (2017).

[47] T. Takeshita, N. C. Rubin, Z. Jiang, E. Lee, R. Babbush, and J. R. McClean, Increasing the representation accuracy of quantum simulations of chemistry without extra quantum resources, Phys. Rev. X **10**, 011004 (2020).

[48] M. Motta, C. Sun, A. T. K. Tan, M. J. O'Rourke, E. Ye, A. J. Minnich, F. G. S. L. Brandão, and G. K.-L. Chan, Determining eigenstates and thermal states on a quantum computer using quantum imaginary time evolution, Nat. Phys. **16**, 205 (2020).

[49] K. Yeter-Aydeniz, R. C. Pooser, and G. Siopsis, Practical quantum computation of chemical and nuclear energy levels using quantum imaginary time evolution and Lanczos algorithms, npj Quantum Inf. **6**, 63 (2020).

[50] W. J. Huggins, J. Lee, U. Baek, B. O'Gorman, and K. B. Whaley, A non-orthogonal variational quantum eigensolver, New J. Phys. **22**, 073009 (2020).

[51] R. M. Parrish and P. L. McMahon, Quantum filter diagonalization: Quantum eigendecomposition without full quantum phase estimation, ArXiv:1909.08925.

[52] J. Cohn, M. Motta, and R. M. Parrish, Quantum filter diagonalization with compressed double-factorized hamiltonians, PRX Quantum **2**, 040352 (2021).

[53] N. H. Stair, R. Huang, and F. A. Evangelista, A multireference quantum Krylov algorithm for strongly correlated electrons, J. Chem. Theory Comput. **16**, 2236 (2020).

[54] C. L. Cortes and S. K. Gray, Quantum krylov subspace algorithms for ground- and excited-state energy estimation, Phys. Rev. A **105**, 022417 (2022).

[55] K. Seki and S. Yunoki, Quantum power method by a superposition of time-evolved states, PRX Quantum **2**, 010333 (2021).

[56] K. Seki and S. Yunoki, Spatial, spin, and charge symmetry projections for a Fermi-Hubbard model on a quantum computer, Phys. Rev. A **105**, 032419 (2022).

[57] K. Kowalski and B. Peng, Quantum simulations employing connected moments expansions, J. Chem. Phys. **153**, 201102 (2020).

[58] B. Peng and K. Kowalski, Variational quantum solver employing the PDS energy functional, Quantum **5**, 473 (2021).

[59] D. Claudino, B. Peng, N. Bauman, K. Kowalski, and T. S. Humble, Improving the accuracy and efficiency of quantum connected moments expansions, Quantum Sci. Technol. **6**, 034012 (2021).

[60] D. Claudino, B. Peng, K. Kowalski, and T. S. Humble, Modeling singlet fission on a quantum computer, J. Phys. Chem. Lett. **14**, 5511 (2023).

[61] W. Kirby, M. Motta, and A. Mezzacapo, Exact and efficient Lanczos method on a quantum computer, Quantum **7**, 1018 (2023).

[62] K. Klymko, C. Mejuto-Zaera, S. J. Cotton, F. Wudarski, M. Urbanek, D. Hait, M. Head-Gordon, K. B. Whaley, J. Moussa, N. Wiebe, W. A. de Jong, and N. M. Tubman, Real-time evolution for ultracompact hamiltonian eigenstates on quantum hardware, PRX Quantum **3**, 020323 (2022).

[63] G. Greene-Diniz, D. Z. Manrique, K. Yamamoto, E. Plekhanov, N. Fitzpatrick, M. Krompiec, R. Sakuma, and D. M. Ramo, Quantum computed Green's functions using a cumulant expansion of the Lanczos method, ArXiv:2309.09685.

[64] A. Hashim, R. K. Naik, A. Morvan, J.-L. Ville, B. Mitchell, J. M. Kreikebaum, M. Davis, E. Smith, C. Iancu, K. P. O'Brien, I. Hincks, J. J. Wallman, J. Emerson, and I. Siddiqi, Randomized compiling for scalable quantum computing on a noisy superconducting quantum processor, Phys. Rev. X **11**, 041039 (2021).

[65] Q. Sun, T. C. Berkelbach, N. S. Blunt, G. H. Booth, S. Guo, Z. Li, J. Liu, J. D. McClain, E. R. Sayfutyarova, S. Sharma, S. Wouters, and G. K. Chan, Pyscf: The python-based simulations of chemistry framework, Wiley Interdiscip. Rev. Comput. Mol. Sci. **8**, e1340 (2017).

[66] Gurobi Optimization, LLC, Gurobi Optimizer Reference Manual (2023), https://www.gurobi.com.

[67] H. Abraham, AduOffei, R. Agarwal, I. Y. Akhalwaya, G. Aleksandrowicz, T. Alexander, M. Amy, E. Arbel, Arijit02, A. Asfaw, *et al.*, Qiskit: An open-source framework for quantum computing (2019).

[68] B. Efron and R. J. Tibishirani, *An Introduction to the Bootstrap* (CRC Press, New York, 1994).

[69] J. R. McClean, S. Boixo, V. N. Smelyanskiy, R. Babbush, and H. Neven, Barren plateaus in quantum neural network training landscapes, Nat. Commun. **9**, 4812 (2018).

[70] J. Chen, H.-P. Cheng, and J. K. Freericks, Flexibility of the factorized form of the unitary coupled cluster Ansatz, J. Chem. Phys. **156**, 044106 (2022).

[71] D. Thouless, Stability conditions and nuclear rotations in the Hartree-Fock theory, Nucl. Phys. **21**, 225 (1960).

[72] I. O. Sokolov, P. K. Barkoutsos, P. J. Ollitrault, D. Greenberg, J. Rice, M. Pistoia, and I. Tavernelli, Quantum orbital-optimized unitary coupled cluster methods in the strongly correlated regime: Can quantum algorithms outperform their classical equivalents?, J. Chem. Phys. **152**, 124107 (2020).

[73] H. R. Grimsley, S. E. Economou, E. Barnes, and N. J. Mayhall, An adaptive variational algorithm for exact molecular simulations on a quantum computer, Nat. Commun. **10**, 3007 (2019).

[74] H. L. Tang, V. O. Shkolnikov, G. S. Barron, H. R. Grimsley, N. J. Mayhall, E. Barnes, and S. E. Economou, Qubit-ADAPT-VQE: An adaptive algorithm for constructing hardware-efficient ansätze on a quantum processor, PRX Quantum **2**, 020310 (2021).

[75] In this work, we take a "logical" qubit to be the information that is stored on a physical qubit (sometimes also known as a "virtual" qubit). This should not be confused with the terminology as used in quantum error correction.

[76] H.-Y. Huang, R. Kueng, and J. Preskill, Predicting many properties of a quantum system from very few measurements, Nat. Phys. **16**, 1050 (2020).

[77] A. Zhao, N. C. Rubin, and A. Miyake, Fermionic partial tomography via classical shadows, Phys. Rev. Lett. **127**, 110504 (2021).

[78] I. D. Kivlichan, J. McClean, N. Wiebe, C. Gidney, A. Aspuru-Guzik, G. K.-L. Chan, and R. Babbush, Quantum simulation of electronic structure with linear depth and connectivity, Phys. Rev. Lett. **120**, 110501 (2018).

[79] G. J. Mooney, Entanglement in superconducting quantum devices and improving quantum circuit compilation, Ph.D. thesis, University of Melbourne, Australia, 2022.

[80] F. B. Maciejewski, Z. Zimborás, and M. Oszmaniec, Mitigation of readout noise in near-term quantum devices by classical post-processing based on detector tomography, Quantum **4**, 257 (2020).

[81] P. D. Nation, H. Kang, N. Sundaresan, and J. M. Gambetta, Scalable mitigation of measurement errors on quantum computers, PRX Quantum **2**, 040326 (2021).

[82] C. Michelot, A finite algorithm for finding the projection of a point onto the canonical simplex of $\mathcal{R}^n$, J. Optim. Theor. Appl. **50**, 195 (1986).

[83] J. A. Smolin, J. M. Gambetta, and G. Smith, Efficient method for computing the maximum-likelihood quantum state from measurements with additive gaussian noise, Phys. Rev. Lett. **108**, 070502 (2012).

[84] X. Bonet-Monroig, R. Sagastizabal, M. Singh, and T. E. O'Brien, Low-cost error mitigation by symmetry verification, Phys. Rev. A **98**, 062339 (2018).

[85] A. J. Coleman, Structure of fermion density matrices, Rev. Mod. Phys. **35**, 668 (1963).

[86] M. J. Knight, H. M. Quiney, and A. M. Martin, Reduced density matrix approach to ultracold few-fermion systems in one dimension, New J. Phys. **24**, 053004 (2022).

[87] D. A. Mazziotti, Pure-N-representability conditions of two-fermion reduced density matrices, Phys. Rev. A **94**, 032516 (2016).

[88] N. C. Rubin, R. Babbush, and J. McClean, Application of fermionic marginal constraints to hybrid quantum algorithms, New J. Phys. **20**, 053020 (2018).

[89] Z. Cai, R. Babbush, S. C. Benjamin, S. Endo, W. J. Huggins, Y. Li, J. R. McClean, and T. E. O'Brien, Quantum error mitigation, Rev. Mod. Phys. **95**, 045005 (2023).

[90] P. Czarnik, A. Arrasmith, P. J. Coles, and L. Cincio, Error mitigation with Clifford quantum-circuit data, Quantum **5**, 592 (2021).

[91] A. M. Dalzell, N. Hunter-Jones, and F. G. S. L. Brandão, Random quantum circuits transform local noise into global white noise, ArXiv:2111.14907.

[92] K. Temme, S. Bravyi, and J. M. Gambetta, Error mitigation for short-depth quantum circuits, Phys. Rev. Lett. **119**, 180509 (2017).