Quantum chemistry as a benchmark for near-term quantum computers

Alexander J. McCaskey1,2,5*, Zachary P. Parks1,2, Jacek Jakowski1,3,5*, Shirley V. Moore1,2, Titus D. Morris1,4, Travis S. Humble1,3 and Raphael C. Pooser1,3,5*

We present a quantum chemistry benchmark for noisy intermediate-scale quantum computers that leverages the variational quantum eigensolver, active-space reduction, a reduced unitary coupled cluster ansatz, and reduced density purification as error mitigation. We demonstrate this benchmark using 4 of the available qubits on the 20-qubit IBM Tokyo and 16-qubit Rigetti Aspen processors via the simulation of alkali metal hydrides (NaH, KH, RbH), with accuracy of the computed ground state energy serving as the primary benchmark metric. We further parameterize this benchmark suite on the trial circuit type, the level of symmetry reduction, and error mitigation strategies. Our results demonstrate the characteristically high noise level present in near-term superconducting hardware, but provide a relevant baseline for future improvement of the underlying hardware, and a means for comparison across near-term hardware types. We also demonstrate how to reduce the noise in post processing with specific error mitigation techniques. Particularly, the adaptation of McWeeny purification of noisy density matrices dramatically improves accuracy of quantum computations, which, along with adjustable active space, significantly extends the range of accessible molecular systems. We demonstrate that for specific benchmark settings and a selected range of problems, the accuracy metric can reach chemical accuracy when computing over the cloud on certain quantum computers.

npj Quantum Information (2019)5:99 ; https://doi.org/10.1038/s41534-019-0209-0

INTRODUCTION
Noisy intermediate-scale quantum (NISQ) devices have been used recently to demonstrate a variety of different small-scale quantum computations.1–5 These demonstrations underscore the progress in developing programmable quantum processing units (QPUs) as well as advances in how to use these devices for scientifically meaningful computations. However, evaluating the impact of these demonstrations is complicated by the interplay of unique features available from different quantum computing technologies (superconducting electronics, trapped ions, etc.) and the multitude of programming choices that influence observed performance. Several properties quantify the intrinsic utility of a QPU including the capacity of the quantum register, the fidelity of the available instructions, the connectivity between register elements, etc. While these hardware-specific parameters characterize progress in quantum hardware development, they do not directly benchmark the performance of a quantum computer for a given computational task.

Benchmarks for application-specific metrics are needed to evaluate the efficiency and applicability of quantum computing for scientific applications. In particular, comparing the benchmark performance of quantum computation to alternative computational approaches should rely on definitions that are meaningful across computational paradigms. Application-specific performance metrics have been used recently to evaluate the utility of near-term QPUs with respect to computational accuracy, including examples from machine learning6 and nuclear physics.7 However, the design and demonstration of an application-specific benchmark to evaluate QPU performance across scalable problem instances has been absent from the literature up to now.

We present a quantum chemistry simulation benchmark to evaluate the performance of quantum computing by defining a series of electronic structure calculation instances that can be realized on current hardware. We outline and investigate an array of parameter choices that influence quantum computational performance. Our approach uses multiple techniques including density matrix purification and active-space reduction via frozen-core approximation and truncation of the virtual space to accommodate hardware limitations such as a limited number of noisy qubits (low circuit width) and a limited achievable circuit depth (the number of parallel gate executions), while staying within the well-known hierarchy of quantum chemistry methods. For calculations on small molecular systems, such as alkali metal hydride molecules with multiple electrons presented here, the complexity of the quantum computation can be reduced to two valence electrons and the equivalent of a hydrogen molecule in minimal basis set. These approximations can be gradually lifted as the quality of quantum devices improves to provide a flexible benchmarking model.

The quantum algorithmic primitive at the center of this benchmark is the state preparation circuit—the unitary coupled cluster (UCC) ansatz7 or hardware-efficient (HWE) ansatz3,8,9 for example—used in the variational quantum eigensolver (VQE).1,10 A noise-resilient algorithm which has been implemented on current noisy QPUs. This hybrid algorithm uses a classical search for the lowest eigenvalue of a given observable using the variational principle and the chosen ansatz, which is prepared and measured on the quantum computer. The utility of this algorithm has been demonstrated in a number of fields, including high-energy and nuclear physics11 in addition to quantum chemistry.9 Its success

1Quantum Computing Institute, Oak Ridge National Laboratory, Oak Ridge, TN 37831, USA. 2Computer Science and Mathematics Division, Oak Ridge National Laboratory, Oak Ridge, TN 37831, USA. 3Computational Sciences and Engineering Division, Oak Ridge National Laboratory, Oak Ridge, TN 37831, USA. 4Physics Division, Oak Ridge National Laboratory, Oak Ridge, TN 37831, USA. 5These authors contributed equally: Alexander J. McCaskey, Jacek Jakowski, Raphael C. Pooser. *email: mccaskeyaj@ornl.gov; jakowskij@ornl.gov; pooserrc@ornl.gov

Published in partnership with The University of New South Wales
is primarily due to the reliance on finite classical computing resources in tandem with invocations of the quantum computer for exponentially scaling aspects of the problem. We use multiple variations of the state-preparation primitives noted above in order to illustrate its effect on performance and accuracy of quantum computations. The benchmark assesses accuracy of VQE for recovering the ground state energy for a series of molecules when using different parameterized trial circuits and error mitigation strategies. The alkali metal hydrides currently included in the benchmark are NaH, KH, and RbH. We implemented calculations of the ground state energies of each of these molecules using four qubits on the 20-qubit IBM Tokyo and 16-qubit Rigetti Aspen superconducting circuit devices.

The variational quantum eigensolver (VQE) has been used previously to calculate ground state energy, and it has become a leading example of quantum-classical computing. Several groups have demonstrated VQE applications using a variety of QPUs and trial circuits to target small molecular systems.\(^ {1, 4, 5, 8, 9, 13-15} \)

VQE recovers the ground state energy of a Hamiltonian by searching for the quantum state that minimizes the observed energy expectation value. The complexity of this classical search through the parameter space depends strongly on the form of the underlying trial circuit. Optimization of the energy expectation value uses preparation and measurement of the parameterized quantum circuit to drive the classical search method executed on a conventional computer. Existing demonstrations of VQE have revealed that despite the algorithm’s noise resilience, error mitigation is required to overcome the extensive noise present in current platforms.\(^ {10-18} \)

For chemistry applications, the main difference between HWE and UCC choices is that the latter provides a rigorous, electron number-conserving procedure for describing a trial wavefunction, but the corresponding circuit has a relatively large depth even for small systems. The HWE ansatz offers shorter circuit depth, but it supports a larger parameter space, and it prepares states with varying numbers of electrons that yield unphysical results.

In ref.\(^ {8} \), the number of parameters, and hence the circuit depth, is reduced for the UCC ansatz by using pre-screening of cluster amplitudes. This pre-screening discards all the excitation operators for which the second order Møller-Plesset perturbation amplitudes are below a chosen threshold. The number of qubits required for the VQE-UCI calculation, as well as the number of parameters required for preparation of the ansatz, are also reduced by using an active space approximation which divides the orbital space into a set of inactive and active orbitals with the occupation of the orbitals in the inactive space remaining unchanged. In the present work we use a frozen core, second-quantized Hamiltonian approximation to further enable mapping of benchmark problems to noisy quantum hardware in a computational framework such as OpenFermion.\(^ {19} \)

Electronic structure calculations require a choice of basis set and simulation approach, which together have an impact on resource cost and accuracy of the simulation. The hierarchy of the theoretical methods and basis sets has been established over the years. The most popular methods in order of increasing accuracy (and computational cost) are Hartree-Fock (HF), density functional theory (DFT), perturbation theories (PT) and configuration interaction (CI), coupled cluster methods, full configuration interaction (FCI). Neither HF nor FCI is used in any practical calculations since the former one is too inaccurate while the latter one is too expensive. Similarly, basis sets of increasing flexibility and size have been developed. The two most popular classes of basis functions are Pople basis sets (for example, 3-21G or 6-31G) and Dunning basis sets (for example cc-pVZ, cc-pVTZ). The smallest available basis set, also called the minimal basis set, is denoted as STO-3G (Slater Type Orbital contracted with 3 Gaussians). Although the STO-3G basis set is not used in practical classical calculations, it represents a good first model for development and benchmarking of new computational methods.

Mapping the quantum chemistry problem to quantum computers involves several steps: (1) molecule specification (structure, charge etc.); (2) generation of the integrals and Hartree-Fock calculations to set an initial reference state; (3) transformation of the Hamiltonian from second quantization formalism to a qubit representation using either the Jordan-Wigner\(^ {20} \) or the Bravyi-Kitaev\(^ {21} \) transformation; (4) generation of a quantum circuit that represents the trial wavefunction; and (5) mapping the quantum circuit to specific quantum computing hardware.

The electronic structure calculations in the benchmark seek the ground state energy configuration of the Hamiltonian described as a sum of a nuclear repulsion \((0\text{-electrons})\) \(H_0\), term, a one-electron term (kinetic energy of electrons plus interaction with core ions) \(H_1\), and a two electron Hamiltonian \(H_2\).

\[
\hat{H} = H_0 + H_1 + H_2
\]

\( \text{where p, q, r, s run over all molecular spin orbitals, } \hat{p}^i, \hat{q} \text{ etc. are corresponding electron creation and annihilation operators, and } H_0 \text{ are matrix elements of the core Hamiltonian (kinetic energy of electrons plus interaction with core ions). The } g_{pq}^i \text{ are two-electron repulsion integrals } g_{pq}^i = \langle p, q | s, r \rangle. \)

The above expression for the Hamiltonian in the chemistry notation can be equivalently written in the so called physicists notation:

\[
\hat{H} = H_0 + \sum_{p,q} h_{pq}^0 \cdot \hat{p}^i \hat{q}^j + \frac{1}{2} \sum_{p,q,r,s} g_{pqrs}^0 \cdot \hat{p}^i \hat{q}^j \hat{r}^k \hat{s}^l
\]

with \(g_{pq}^0\) being an anti-symmetrized repulsion integral

\[
g_{pq}^0 = g_{pq}^0 - g_{qp}^0 = \langle p, q | s, r \rangle - \langle q, p | r, s \rangle.
\]

To reduce depth and enable execution on noisy near-term quantum computers, we neglect the contribution to electronic correlation from the lowest energy core electrons. We treat only the interaction of correlated electrons with core electrons in an average mean-field fashion. Thus we generate the following reduced effective Hamiltonian:

\[
\hat{H} = H_0 + H_1 + H_2
\]

with \(H_0\) given by

\[
H_0 = E_{\text{nucl}} + \sum_a \left( h_a^0 + \frac{1}{2} \sum_b g_{ab}^0 \right)
\]

where \(a\) and \(b\) run over frozen-core spin orbitals. Similarly, the one-body term is given by

\[
H_1 = \sum_{p,q} \hat{p}^i \hat{q}^j \left( h_p^q + \frac{1}{2} \sum_a g_{pq}^a \right)
\]

and the two-body part is

\[
H_2 = \frac{1}{4} \sum_{p,q,r,s} g_{pqrs}^0 \cdot \hat{p}^i \hat{q}^j \hat{r}^k \hat{s}^l.
\]

Here we note that in the last three expressions, the indices \(p, q, r, s\) run over active-spin orbitals instead of all spin orbitals. Indices \(a\) and \(b\) run over frozen-core orbitals which are not active. Finally, one can exclude the specific virtual orbitals from the active space. Figure 1 illustrates partitioning of orbital space onto frozen core, active space and inactive virtual space.

In order to calculate the ground state energy of Eq. (1) using VQE, one begins by mapping the fermionic representation to an
equivalent spin-based Hamiltonian
\[
\hat{H} = \sum_i h_i^e \sigma_i^e + \sum_{i,j,a} h_{ij}^\alpha \sigma_i^\alpha \sigma_j^\alpha + \sum_{i,j,k,a,b} h_{ijk}^\alpha \sigma_i^\alpha \sigma_j^\beta \sigma_k^\gamma + \ldots
\]
via well-known transformations. Here, Pauli tensor product \(\sigma_i^\alpha \in \{\sigma_x, \sigma_y, \sigma_z, I\}\) acts on the \(i\)th qubit. Next, the qubit register is initialized to an unentangled reference state \(|\psi_0\rangle\), followed by the application of a parameterized unitary \(U(\theta)\) producing the state \(|\psi(\theta)\rangle = U(\theta)|\psi_0\rangle\). Measurements dictated by the Pauli tensor product terms in Eq. (8) are then performed on this state to compute the individual Pauli term expectation values, which are then contracted with term coefficients to produce the variational energy \(E(\theta) = \langle \psi(\theta)|\hat{H}|\psi(\theta)\rangle\). The parameters \(\theta\) are then updated as part of a user-specified classical optimization routine and the entire process is repeated until convergence to the minimum of \(E(\theta)\).

RESULTS

Our benchmark performs calculations of ground state energy for three different alkali metal hydrides—NaH, KH, and RbH. For each of these molecules, we restricted the active space by freezing inner-core electrons, leaving four active orbitals with two fermionic degrees of freedom. All fermionic Hamiltonians were mapped to spin Hamiltonians via the Jordan-Wigner transform.

To improve the accuracy of our results, we employed a new error mitigation technique called reduced density matrix (RDM) purification (see the Methods section) while running the benchmarks on both the IBM and Rigetti quantum computers. Figure 3 shows the results of RDM purification for NaH on Tokyo (top) and Aspen (bottom) using the ucc-1 ansatz. Using this error mitigation approach we observe chemical accuracy at the optimal angle on the Rigetti QPU. We also considered a more algorithmically correct 3-parameter extension to the UCC ansatz (labeled ucc-3; see Fig. 7 in the Methods section) for NaH. We classically optimized the parameters via the COBYLA method as shown for IBM (top) and Rigetti QPUs (bottom) in Fig. 4. These computations employed 1000 shots on Rigetti and 8192 shots on IBM. The inset plots demonstrate the difference between purified energy points and the FCI energy, with the solid line denoting chemical accuracy (0.0016 Ha). The Rigetti results obtain chemical accuracy for this three parameter optimization. This marks a key milestone in NISQ computing, where advanced error mitigation techniques have enabled chemical accuracy over remote cloud quantum computing resources on a true variational optimization calculation.

Finally, we applied RDM purification to the computation of the NaH energy surface. We chose four inter-atomic distance values, \(R\), and computed \(E(R)\) with the one-parameter UCC ansatz. We repeated this five times for each value of \(R\) to gather statistics. Figure 5 shows the results of this computation, with errors on the order of 0.01 (smaller than the markers).

DISCUSSION

Near-term quantum computation is limited in the total number of operations that can be performed. It is limited in the number of two-qubit entangling gates that can be present in a program.
VQE algorithm, an active space reduction of the electronic—cases must take these limitations into account. A J. McCaskey et al.

parameter $\theta$ that enables cross-platform comparison of various scientific fields. Further, the basic form of the ansatz used here apply to other nuclear physics effective field theory computations and in quantum field theory, for example. This means that in addition to quantum chemistry, if a quantum computer performs well at this task as the size of the algorithm scales up, then it will be able to achieve a quantum advantage. However, error mitigation as it exists today is not yet scalable. This means that the more error mitigation required to reach a good accuracy, the less scalable the algorithm is on that machine. Not surprisingly, none of the publicly available hardware is capable of a quantum advantage, but this does not preclude future iterations of the same hardware from reaching this regime.

Finally, we note that while we ran the benchmark on two different hardware platforms, we do not take the performance difference between the two to be indicative of long-term trends or to be definitive enough to declare a "winner". Rather, we intend for the benchmark to provide a user with information on what to expect when running a given application on a particular machine, independent of other devices.
Table 1. Computed ground state energies

| Molecule | Ansatz | Tapering | QPU   | EM   | $E$         | $\Delta E$ |
|----------|--------|----------|-------|------|-------------|------------|
| NaH      | ucc-1  | No       | Tokyo | none | $-160.122 \pm 0.002$ | 0.181      |
| NaH      | ucc-1  | No       | Tokyo | ro+i-re | $-160.347 \pm 0.032$ | -0.044     |
| NaH      | ucc-1  | No       | Aspen | none | $-159.980 \pm 0.004$ | 0.323      |
| NaH      | ucc-1  | No       | Tokyo | none | $-159.973 \pm 0.004$ | 0.330      |
| NaH      | ucc-1  | No       | Tokyo | rdm  | $-160.279 \pm 0.004$ | 0.024      |
| NaH      | ucc-1  | No       | Aspen | none | $-159.917 \pm 0.013$ | 0.39       |
| NaH      | ucc-1  | No       | Tokyo | rdm  | $-160.301 \pm 0.004$ | 0.002      |
| NaH      | ucc-1  | No       | Aspen | none | $-160.049 \pm 0.005$ | 0.254      |
| NaH      | ucc-1  | No       | Tokyo | rdm  | $-160.297 \pm 0.005$ | 0.006      |
| NaH      | hwe    | Yes      | Tokyo | none | $-160.263 \pm 0.009$ | 0.040      |
| NaH      | hwe    | Yes      | Tokyo | ro+i-re | $-160.287 \pm 0.002$ | 0.016      |
| NaH      | hwe    | No       | Tokyo | none | $-160.037 \pm 0.005$ | 0.266      |
| NaH      | hwe    | No       | Tokyo | ro+i-re | $-160.085 \pm 0.007$ | 0.218      |
| KH       | ucc-1  | No       | Tokyo | ro+i-re | $-593.434 \pm 0.002$ | 0.141      |
| KH       | ucc-1  | No       | Tokyo | ro+i-re | $-593.559 \pm 0.035$ | 0.016      |
| RbH      | ucc-1  | No       | Tokyo | none | $-2908.058 \pm 0.001$ | 0.067      |
| RbH      | ucc-1  | No       | Tokyo | ro+i-re | $-2908.158 \pm 0.004$ | -0.033     |

This table details the computed energies ($E$) on the 20-qubit IBM Tokyo and the 16-qubit Rigetti Aspen QPUs for various alkali metal hydrides, ansätze, and error mitigation strategies leveraged (EM-none, ro+i-re for readout-error and Richardson extrapolation, or rdm for RDM purification). The FCI energies for 4-qubit frozen-core NaH, RbH, and KH are $-160.34597$, $-2908.125112$, and $-593.5747682$, respectively.

Fig. 3  The dependence of the energy as a function of $\theta$ for the ucc-1 ansatz for NaH with and without RDM purification for IBM Tokyo (top) and Rigetti QCS (bottom).

Fig. 4  Energy as a function of the optimization iteration for the 3-parameter UCC (ucc-3) ansatz before and after RDM purification on the IBM Tokyo (top) and Rigetti QCS (bottom). The inset plots demonstrate the distance of the computed energy to the FCI energy, with the solid black line denoting chemical accuracy.
that for systems, such as alkali hydrides within frozen-core approximation. Note \( O_{\text{UCCSD}} \) circuit scales as \( O_{\text{UCCSD}} \) ansatz is sufficient for simulation via the application of discrete frozen-core approximation, we also reduce the number of qubits necessary for high accuracy, which is used as a benchmark metric.

Our approach first seeks to reduce the complexity of the Hamiltonian in an effort to optimize the width of our quantum circuits. In addition to the frozen-core approximation, we also reduce the number of qubits necessary for simulation via the application of discrete \( Z_2 \) symmetries.\(^{24}\) Next, we choose a particular trial unitary class. Finally, we consider our benchmarking results as a function of various error mitigation post-processing techniques.

### Trial circuits

The choice of ansatz dictates the depth of the program executed and therefore the level of noise present in the computation. The number of parameters in the circuit determines the difficulty of the classical optimization step and the number of individual QPU calls required.

In the UCC method, the wavefunction is represented using the exponentiated operator\(^{1,2,26}\)

\[
|\psi(\theta)\rangle = e^{T^\dagger} |\psi_0\rangle
\]

where \( |\psi_0\rangle \) is a reference state; for example, the Hartree-Fock solution. The symbol \( T \) (\( T^\dagger \)) is a particle excitation (de-excitation) operator, given by

\[
T = \sum_{k=1}^{M} T_k(\theta)
\]

\[
T_k(\theta) = \sum_{\alpha<\beta} \theta_{\alpha\beta} a_\alpha^\dagger a_\beta
\]

\[
T_1(\theta) = \frac{1}{4} \sum_{\alpha<\beta<\gamma<\delta} \theta_{\alpha\beta\gamma\delta} a_\alpha^\dagger a_\beta^\dagger a_\gamma a_\delta
\]

where \( M \) is the number of electrons and the \( \theta \) parameters map directly to the parameters found in the circuit decomposition of this unitary operator. For UCC single doubles (UCCSD), the excitation operator is truncated to single and double excitations only, \( T = T_1(\theta) + T_2(\theta) \). In general, the UCCSD ansatz is sufficient to map the exact FCI solution for 2-electron systems, such as alkali hydrides within frozen-core approximation. Note that for \( N \) spin-orbitals and \( M \) electrons, the number of parameters for the UCCSD circuit scales as \( O(N^2 M^3) \),\(^8\) and therefore the depth for even modest sized problems grows well past what is implementable on currently available quantum computers. For a four-qubit, 2-electron problem, the depth of this circuit is on the order of 100, and the number of instructions is about 150. Thus, symmetry reduction is required to implement the ansatz on current hardware.

One such reduction uses the fact that the spin terms that make up the exponential argument of Eq. (9) for a hydrogenic problem (four orbitals, two fermions) consist of eight four-site terms that all act identically on the Hartree-Fock state. Therefore we can approximate the state with only one term. Here, we use \( V_{\theta} X_1 X_2 X_3 \) so that our parameterized unitary becomes

\[
U(\theta) = e^{iV_{\theta} X_1 X_2 X_3},
\]

The circuit for this unitary is shown in Fig. 6. This circuit structure is a computational primitive that recurs in larger systems. This CX ladder structure is therefore a prime example of an early benchmark for quantum chemistry on quantum computers—if the quantum device cannot adequately execute this circuit, it will perform poorly on all larger fermionic systems. On the other hand, one can map larger molecules onto this hydrogenic system via a frozen-core approximation, and apply this circuit primitive to produce energy metrics for a cross-platform comparison.

We note that this ansatz cannot produce FCI energy for four-qubit, 2-electron problems, but it provides a UCCSD solution. That is the single parameter \( \theta \) controls the contribution from doubly excited determinants:

\[
\psi(\theta) = \frac{1}{\sqrt{2}} \left( |1010\rangle + c(\theta)|0101\rangle \right),
\]

where \( |1010\rangle \) is a reference ground state Slater determinant, \( |0101\rangle \) is a doubly excited Slater determinant and \( N \) is normalization factor. To correct for this, we append circuits that address additional two-qubit subsystems in order to expand the total accessible Hilbert space represented by the quantum register (see Fig. 7). This extended UCC ansatz introduces an additional parameter that controls a single excitation for each two-qubit spin subspace, allowing for the electronic excitations to spread to fully cover all orbitals considered active in the Hamiltonian.

For completeness, we also consider an implementation of the hardware-efficient ansatz in ref.\(^3\). The circuit is composed of alternating layers of single-qubit rotations and two-qubit entangling operations (based on the connectivity structure of the hardware). The number of parameters in this circuit grows as \( N(3d+2) \), where \( N \) is the number of qubits and \( d \) is the number of rotation-\( d \)-entangling layers. This ansatz is more amenable to near-term execution than a naive implementation of the UCC ansatz due to the controlled growth in its depth. However, the increase in the number of variational parameters leads to an increase in costly QPU executions (which are remote network calls for most available NISQ devices). Furthermore, this ansatz has been shown to introduce barren plateaus—regions where the probability that the gradient in a given direction is non-zero becomes exponentially small as the number of qubits increases.\(^7\)

### Error mitigation strategies

Executing quantum circuits on current NISQ hardware will produce results that do not compare well with classical theoretical values. This is due to a number of factors, including intrinsic systematic noise present during execution and qubit measurement readout errors. Therefore, in order to produce valid results, some form of error mitigation must be employed. We consider the type of and quantity of required error mitigation strategies as a dimension of our near-term, quantum benchmarking approach. We note that required strategies may vary from one QPU to another, and therefore dictate which QPU may be more well-suited for the given application or task. It is therefore important to parameterize our benchmark with respect to these strategies. In this work, we examine three such mitigation strategies: readout-error mitigation, entangling gate error rate extrapolation, and reduced density matrix purification.
We used a local, spatially uncorrelated readout-error model that requires information about the probability of an unexpected bit flip during qubit measurement. For each qubit, we compute the probability that a $|1\rangle$ was observed when a $|0\rangle$ was expected, and the probability a $|0\rangle$ was observed when a $|1\rangle$ was expected, $p_{i}(0|1), p_{i}(1|0)$, respectively. Then, our readout-error corrected expectation values are computed from the experimentally observed bit string counts as

$$
\langle Z\ldots Z \rangle = \sum_{i} \omega_{i} \prod_{j \in \text{sites}(Z \ldots Z)} \left[ \frac{(-1)^{j} - p_{j}}{1 - p_{j}} \right],
$$

where $p(x)$ is the probability of seeing bit string $x$ and $p_{j} = p_{j}(0|1) \pm p_{j}(1|0)$. $i$ sites($Z\ldots Z$) represents the qubit indices represented in the given measured Pauli term.

To mitigate against systematic two-qubit entangling gate noise, we implemented the zero-noise extrapolation technique put forth in ref.19. We assume a generic two-qubit white noise error channel $\mathcal{N} = (1 - e) \mathbb{I} + e/4$. We emulat increasing $\epsilon$ by introducing pairs of CNOT gates, serving as a noisy identity. We introduce $r$ pairs of these entangling operations, compute the energy produced by the circuit each time, and extrapolate back to $r = 0$ to produce the noiseless energy.

Finally, we adapted a McWeeny purification scheme of non-idempotent density matrices. We implemented a mixed-state purification approach that depends on the computation of the two-body reduced density matrix (RDM). Typically, VQE seeks the expectation value of an operator that can be expressed as a sum of individual, weighted Pauli tensor products. If we instead stay in the fermionic framework, we seek the minimum of the following energy expression

$$
E(\theta) = \langle \psi(\theta)| H |\psi(\theta) \rangle
= \sum_{\alpha \beta} h_{\alpha \beta} \langle \psi(\theta) | a_{\alpha}^{\dagger} a_{\beta} | \psi(\theta) \rangle + \frac{1}{2} \sum_{p q r s} h_{pqrs} \langle \psi(\theta) | a_{p}^{\dagger} a_{q}^{\dagger} a_{r} a_{s} | \psi(\theta) \rangle,
$$

where $\langle \psi(\theta) | a_{\alpha}^{\dagger} a_{\beta} | \psi(\theta) \rangle$ is the two-body RDM (2-RDM), $\rho_{pqrs}$. From the study of n-representability theory, an appropriate trace of the 2-RDM yields the one-body RDM. Therefore the 2-RDM is sufficient for computing the overall energy of the system.29 We can compute these 2-RDM elements by constructing all physically-relevant $a_{\alpha}^{\dagger} a_{\beta} a_{\gamma} a_{\delta}$ operators, mapping to the spin representation, executing the chosen parameterized ansatz, and performing measurements dictated by these transformed spin terms. This computation, when done in the presence of systematic noise, produces a statistical mixture of many pure states. Recall that the variational principle ensures that $E_{\theta} \leq E(\theta)$ for some pure state with corresponding eigenvalue $E_{\theta}$. We therefore employ a strategy for purifying these RDM elements, following the well-known McWeeny purification formula.30 To improve our estimate of $\rho_{pqrs}$ and therefore mitigate against inherent error on the hardware, we purify the 2-RDM via the iterative approach $\rho_{pqrs} \rightarrow \frac{3(\rho_{pqrs})^{2} - 2(\rho_{pqrs})^{3}}{\rho_{pqrs}^{2} - 2(\rho_{pqrs})^{3}}$ until $\text{Tr}(\rho_{pqrs} - \rho_{pqrs}^{2}) < \epsilon$, for $\epsilon \ll 1$. Note that here tensor multiplication is defined as the trace $C_{pqrs} = A_{pq}^{\dagger} B_{rs}^{\dagger}$ (Einstein summation implied). We note for clarity that this procedure is only appropriate for 2 electron systems, and will not provide guidance for error mitigation with more than 2 electron simulations. In order to purify a 2-RDM encountered in systems with more than 2 electrons, one must utilize the semidefinite constraints provided by N-representability, as in ref.29. For our current purposes, this was beyond the need and scope of this work. In the future, we plan to extend this benchmark with error mitigation beyond 2 electron systems, which will require an extension to the purification scheme.

In order to estimate the statistical error of the purified energies, we utilized a long established process called bootstrapping.31 This consists of resampling each measured circuit according to probabilities estimated from the measurement results. We used the resampled circuit outcomes to produce a 2-RDM, which then was used to derive raw and purified energies. This process was repeated to create 10,000 bootstrapped pairs of raw and purified energies for each experiment, which provided an estimate of the the mean and standard deviation. As a check, we verified that the bootstrapped error estimates were almost identical to those produced by the variance estimate for non purified energies also used in this work, which indicates it is likely a good measure of the error for the purification process as well.

Software implementation

Enabling a high-level, application-based benchmarking capability for near-term quantum computation requires an extensible and modular approach that abstracts away the underlying hardware and the benchmark application domain. The goal is to provide an executable that can be quickly and easily installed and an input deck that is expressive and enables one to tailor all available benchmark parameters. Our approach provides a hardware-independent benchmark that can be extended to new domains for application-centric and algorithmic primitive benchmarking. We have extended the Eclipse XACC quantum-classical programming framework32 and provided a Pythonic approach for designing and executing benchmarks across the major available QPUs. Our approach extends the Python API provided by XACC with support for runtime-contributed service interfaces for the various aspects of application-centric and primitive benchmarking. The benchmark takes an INI file as input. This file describes the benchmark to be run, including the algorithm to use (VQE), input pertinent to the algorithm (trial circuit, optimizer, etc.), molecular integral generation routine, error mitigation strategies to implement (XACC automates the error mitigation strategies described in the previous section), and the target QPU to benchmark. A typical input file is shown in Listing 1. Providing a cross-platform benchmark capability is therefore a matter of distributing these input files for execution on currently available QPUs. Further details on the software framework are given elsewhere.33

Listing 1. A typical benchmark input file for NaH molecule with STO-3G basis set. Notice that indexes 0–9 in the spin orbitals lists correspond spin-up orbitals whereas 10–19 correspond to spin-down orbitals.

[XACC]
accelerator: ibm:ibmq_20_tokyo
algorithm: vqe
[Error Mitigation]
readout-error: True
richardson-extrapolation: True
[VQE]
optimizer: cobyla
[Ansatz]
name: ucc-3
[Molecule]
basis: stoe-3g
group: Na 0.0 0.0 0.0
H 0.0 0.0 0.1 9.14388
frozen-spin-orbitals: [0,1,2,3,4,10,11,12,13,14]
active-spin-orbitals: [5,9,15,19]

DATA AVAILABILITY

The data that support the findings of this study are available from the authors upon request.

CODE AVAILABILITY

The code used to generate the data in this paper is available as part of the open source XACC programming environment at https://github.com/eclipse/xacc.
REFERENCES

1. O’Malley, P. J. J. et al. Scalable quantum simulation of molecular energies. Phys. Rev. X 6, 031007 (2016).
2. Linke, N. M. et al. Experimental comparison of two quantum computing architectures. Proc. Natl Acad. Sci. USA 114, 3305–3310 (2017).
3. Kandala, A. et al. Hardware-efficient variational quantum eigensolver for small molecules and quantum magnets. Nature 549, 242–246 (2017).
4. Dumitrescu, E. F. et al. Cloud quantum computing of an atomic nucleus. Phys. Rev. Lett. 120, 210501 (2018).
5. Klco, N. & Savage, M. Digitization of scalar fields for quantum computing. Phys. Rev. A 99, 052335 (2019).
6. Hamilton, K. E., Dumitrescu, E. F. & Pooser, R. C. Generative model benchmarks for superconducting qubits. Phys. Rev. A 99, 052335 (2019).
7. Shen, Y. et al. Quantum implementation of the unitary coupled cluster for simulating molecular electronic structure. Phys. Rev. A 95, 020501 (2017).
8. Romero, J. et al. Strategies for quantum computing molecular energies using the unitary coupled cluster ansatz. Quant. Sci. Technol. 4, 014008 (2018).
9. Ryabinkin, I. G., Yen, T.-C., Genin, S. N. & lamoyav, A. F. Qubit coupled cluster method: a systematic approach to quantum chemistry on a quantum computer. J. Chem. Theory. Comput. 14, 6317–6326 (2018).
10. Aspuru-Guzik, A., Dutot, A. D., Love, P. J. & Head-Gordon, M. Simulated quantum computation of molecular energies. Science 309, 1704–1707 (2005).
11. Klco, N. et al. Quantum-classical computation of Schwinger model dynamics using quantum computers. Phys. Rev. A 98, 023311 (2018).
12. Peruzzo, A. et al. A variational eigenvalue solver on a photonic quantum processor. Nat. Commun. 5, 4213 (2014).
13. Kandala, A. et al. Extending the computational reach of a noisy superconducting quantum processor. arXiv e-prints arXiv:1805.04492 (2018).
14. Hempel, C. et al. Quantum chemistry calculations on a trapped-ion quantum simulator. Phys. Rev. X 8, 031022 (2018).
15. Shen, Y. et al. Quantum implementation of the unitary coupled cluster for simulating molecular electronic structure. Phys. Rev. A 95, 020501 (2017).
16. Temme, K., Bravyi, S. & Gambetta, J. M. Error mitigation for short-depth quantum circuits. Phys. Rev. Lett. 119, 180509 (2017).
17. Bonet-Monroig, X., Sagastizabal, R., Singh, M. & O’Brien, T. E. Low-cost error mitigation by symmetry verification. Phys. Rev. A 98, 062339 (2018).
18. Li, Y. & Benjamin, S. C. Efficient variational quantum simulator incorporating active error minimization. Phys. Rev. X 7, 021050 (2017).
19. McClean, J. R. et al. OpenFermion: the electronic structure package for quantum computers. arXiv e-prints arXiv:1710.07629 (2017).
20. Jordan, P. & Wigner, E. Über das Paulische Äquivalenzverbot. Zeitschrift für Physik 47, 631–651 (1928).
21. Bravyi, S. & Kitaev, A. Y. Fermionic quantum computation. Annal. Phys. 298, 210–226 (2002).
22. Seeley, J. T., Richard, M. J. & Love, P. J. The bravyi-kitaev transformation for quantum computation of electronic structure. J. Chem. Phys. 137, 224109 (2012).
23. Powell, M. J. D. Direct search algorithms for optimization. Acta Numerica 7, 287 (1998).
24. Bravyi, S., Gambetta, J.M., Mezzacapo, A. & Temme, K. Tapering off qubits to simulate fermionic Hamiltonians. arXiv e-prints arXiv:1701.08213 (2017).
25. Boixo, S. et al. Characterizing quantum supremacy in near-term devices. Nat. Phys. 14, 595 (2018).
26. Taube, A. G. & Bartlett, R. J. New perspectives on unitary coupled-cluster theory. Int. J. Quantum Chem. 106, 3393–3401 (2006).
27. McClean, J. R., Boixo, S., Smelyanskiy, V. N., Babbush, R. & Neven, H. Barren plateaus in quantum neural network training landscapes. Nat. Commun. 9, 4812 (2018).
28. Morris, T. Improved optimization for effective field theory simulations with a ucc ansatz (in preparation).
29. Rubin, N. C., Babbush, R. & McClean, J. Application of fermionic marginal constraints to hybrid quantum algorithms. New J. Phys. 20, 053020 (2018).
30. Truflander, L. A., Dianzoga, R. M. & Bowler, D. R. Communication: generalized canonical purification for density matrix minimization. J. Chem. Phys. 144, 091102 (2016).
31. Efron, B. Bootstrap methods: another look at the jackknife. Annal. Statistics 7, 1 (1979).
32. McCaskey, A. et al. A language and hardware independent approach to quantum-classical computing. Softw. X 7, 245–254 (2018).
33. McCaskey, A. & Parks, Z. A dynamic service-oriented platform for benchmarking quantum computers (In preparation).

ACKNOWLEDGEMENTS

This manuscript has been authored by UT-Battelle, LLC under Contract No. DE-AC05-00OR22725 with the U.S. Department of Energy. The authors acknowledge fruitful discussions with E. Dumitrescu. R.C.P. acknowledges helpful discussions with R.S. Bennink. The authors acknowledge DOE ASCR funding under the Quantum Computing Testbed Pathfinder program, FWP number ERKJ332. Z.P. was supported in part by an appointment to the Oak Ridge National Laboratory HERE Program, sponsored by the U.S. Department of Energy and administered by the Oak Ridge Institute for Science and Education. This research used quantum computing system resources supported by the U.S. Department of Energy, Office of Science, Office of Advanced Scientific Computing Research program office.

AUTHOR CONTRIBUTIONS

R.C.P. conceived of the project, supervised the research, and interpreted results. A.J.M. and Z.P. conducted the experiments by developing the benchmarking software, programming the quantum computers, and taking data. S.V.M. contributed to design of the experiments and interpreted results. J.J. contributed quantum chemistry Hamiltonians and developed custom ansatz for optimization. T.D.M. implemented reduced density matrix purification and bootstrapping. T.S.H. interpreted results. All authors participated in writing the paper.

COMPETING INTERESTS

The authors declare no competing interests.

ADDITIONAL INFORMATION

Correspondence and requests for materials should be addressed to A.J.M., J.J. or R.C.P.

Reprints and permission information is available at http://www.nature.com/reprints

Publisher’s note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made. The images or other third party materials in this article are included in the article’s Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the article’s Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this license, visit http://creativecommons.org/licenses/by/4.0/.

This is a U.S. government work and not under copyright protection in the U.S.; foreign copyright protection may apply 2019