Preserving Symmetries for Variational Quantum Eigensolvers in the Presence of Noise

George S. Barron, Bryan T. Gard, Orien J. Altman, Nicholas J. Mayhall, Edwin Barnes, and Sophia E. Economou

Phys. Rev. Applied **16**, 034003 — Published 1 September 2021

DOI: 10.1103/PhysRevApplied.16.034003

Preserving Symmetries for Variational Quantum Eigensolvers in the Presence of Noise

George S. Barron,[1,*] Bryan T. Gard,[1] Orien J. Altman,[1]
Nicholas J. Mayhall,[2] Edwin Barnes,[1] and Sophia E. Economou[1]

[1] *Department of Physics, Virginia Tech, Blacksburg, VA 24061, U.S.A*
[2] *Department of Chemistry, Virginia Tech, Blacksburg, VA 24061, U.S.A*

(Dated: August 16, 2021)

One of the most promising applications of noisy intermediate-scale quantum computers is the simulation of molecular Hamiltonians using the variational quantum eigensolver. We show that encoding symmetries of the simulated Hamiltonian in the VQE ansatz reduces both classical and quantum resources compared to other, widely available ansätze. Through simulations of the H_2 molecule and of a Heisenberg model on a two-dimensional lattice, we verify that these improvements persist in the presence of noise. This is done using both real IBM devices and classical simulations. We also demonstrate how these techniques can be used to find molecular excited states of various symmetries using a noisy processor. We use error mitigation techniques to further improve the quality of our results.

I. INTRODUCTION

Quantum computers are believed to be one of the most promising technologies currently being developed that will help extend the reach of scientific discovery. This may be achieved through quantum simulation [1], which leverages the properties of a quantum processing unit (QPU) to simulate naturally occurring quantum mechanical systems. One of the most popular algorithms within the reach of near-term devices is the variational quantum eigensolver (VQE) [2–7].

This algorithm falls under a more general class of algorithms known as hybrid variational quantum algorithms [8–14]. The general principle of these algorithms is to use a feedback loop between the quantum and classical computers to minimize a predefined cost function. This method has been applied to a variety of quantum systems in both theoretical [15–27], and experimental [2, 3, 5, 12, 28–31] contexts. In the case of the VQE, the predefined function is the expectation value of the simulated Hamiltonian with respect to the state of the QPU. Additionally, a variety of techniques are available that use the VQE to find higher excited states of such systems [6, 32, 33]. These variational algorithms in general are of particular interest for near-term devices due to their proposed noise resilience. Notably, Ref. [4] has demonstrated noise resilience against coherent errors, and Ref. [34] proved noise resilience against incoherent errors for a special class of variational algorithms.

The advantage of hybrid variational algorithms lies in their ability to exploit advanced classical computational resources without having to store the wavefunction on a classical computer. Wavefunctions are instead prepared and measured on the QPU. The success of these algorithms depends sensitively on challenges involving optimizing rapidly varying functions with many parameters, limited computational resources, and the presence of noise [24, 35–45]. Work has been done to manipulate the input and output of the QPU so as to mitigate and characterize the error [46–49]. Another, perhaps more fundamental issue is related to the complexity of the quantum circuits used to prepare the wavefunctions. This complexity in turn is fundamentally related to the number of parameters needed to describe arbitrary states. Hence, a fundamental question is "how likely is it that a given parameterized quantum circuit (a variational ansatz) will represent the targeted state with sufficient accuracy?" Typically there are no accuracy guarantees for a given ansatz. One approach to designing ansätze with tunable accuracy is to iteratively build a customized variational ansatz in a "problem tailored" fashion, as proposed with the adaptive derivative assembled problem-tailored VQE (ADAPT-VQE) algorithm [50, 51]. This is accomplished by growing the ansatz operator-by-operator, each time selecting the operator which creates the largest change to the objective function, using the VQE as a subroutine at each step. This is in contrast to alternative approaches that fix the number of parameters from the start. In Ref. [22], we showed how to create quantum circuits that have exactly the necessary number of variational parameters to describe any state in the relevant symmetry subspace of the Hilbert space.

One of the advantages of Ref. [22] is that the ansatz is guaranteed to find the correct state in the absence of noise, both for ground and for excited states. Other algorithms that accomplish excited state simulation are sometimes iterative and rely on previous results of a VQE, thus introducing computational overheads. Despite these advantages of symmetry-enforcing circuits, however, it is not clear how the presence of noise might affect their performance. The operators which create noise clearly will not commute with the same symmetries as the target Hamiltonian, and as such, might drive the system into undesirable symmetry subspaces. Will a circuit that preserves these symmetries perform worse due to a lack of relevant error mitigating degrees of freedom?

In this work, we consider the effect that noise has on

* gbarron@vt.edu

the ability of a VQE to preserve symmetries. The simulations here are done using IBM's Qiskit [52] software, using noise models derived from real superconducting devices. We also test our methods on real IBM devices. We perform these simulations and experiments using our previously developed techniques [22] for encoding symmetries of the problem Hamiltonians directly into the state preparation circuit. We test our approach on H_2 and on a two-dimensional Heisenberg model, finding in both cases that our ansätze can outperform standard, ad hoc ansätze. We adapt this approach to solving for excited states that obey certain symmetries. This allows us to guarantee the ability of the ansatz to find the desired state, and in fact, accomplish this with a minimal number of variational parameters. Additionally, our approach for producing excited states does not rely on previous runs of VQEs and is parallelizable for the different excited states calculated.

This paper is structured as follows. In Sec. II we provide the necessary background to make this paper self-contained. Specifically, in Sec. II A, we review the theory behind VQE and its application to quantum chemistry problems. In Sec. II B we review our earlier work on symmetry preserving VQE ansätze. In Secs. III and IV, we perform simulations on the H_2 molecule and analyze the performance of these algorithms in noiseless and noisy simulation contexts, respectively. In Sec. IV we use, benchmark, and compare a variety of error mitigation techniques, as well as discuss challenges associated with noisy optimization. In Sec. V we compare the ability of the symmetry preserving and ad hoc ansätze to preserve desired symmetries with both noisy simulations and experiments on IBM devices. In Sec. VI we present the results of this procedure for excited state simulation. In Sec. VII we discuss potential future work incorporating device connectivity. In Sec. VIII, we give a summary of our conclusions.

II. BACKGROUND

A. Variational Quantum Eigensolver

The goal of a VQE is to minimize an objective function of the form $f(\vec{\theta}) = \left\langle \psi(\vec{\theta}) \right| H \left| \psi(\vec{\theta}) \right\rangle$, where $\left|\psi(\vec{\theta})\right\rangle = U(\vec{\theta})\left|\psi_0\right\rangle$ is the state of the QPU after initializing the register to some initial state $|\psi_0\rangle$ and applying a circuit with unitary $U(\vec{\theta})$, where $\vec{\theta}$ is a vector of variational parameters chosen by an optimization algorithm. Per the variational principle, $f(\vec{\theta}) \geq E_0$ where E_0 is the ground state energy of the Hamiltonian H.

We primarily focus on the second quantized molecular Hamiltonian

$$\hat{H}_f = \sum_{ij} h_{ij} \hat{a}_i^\dagger \hat{a}_j + \sum_{ijkl} g_{ijkl} \hat{a}_i^\dagger \hat{a}_j^\dagger \hat{a}_k \hat{a}_l, \tag{1}$$

where $\hat{a}$ ($\hat{a}^\dagger$) is the fermionic annihilation (creation) operator. The quantities h_{ij} and g_{ijkl} are the single and double electron integrals respectively, which can be computed efficiently on a classical computer with existing software. For these calculations we use the STO-3G basis set. To evaluate the objective function $f(\vec{\theta})$ on the QPU, one must map the fermionic operators to a set of qubit operators in a way that accounts for the fermionic anti-commutation relations.

This can be accomplished with several different methods, including the Jordan-Wigner (JW) [53], parity [54], Bravyi-Kitaev (BK) [55], or Bravyi-Kitaev Super Fast (BKSF) [56] mappings. In this work we choose the JW mapping as it directly maps occupations of fermionic spin orbitals to excitations of qubits. The result is that our ansatz consists of only nearest-neighbor two-qubit gates, which we choose to preserve particle number and spin projection. This choice is not necessary, but other maps would require gates that act on more than two qubits to attain the same effect. To achieve this, one would need to find mapping-specific gates that preserve the desired expectation values of the states on which they act, but this is beyond the scope of the current work.

The resulting Hamiltonian appearing in the objective function $f(\vec{\theta})$ is a weighted sum of Pauli strings $\hat{H} = \sum_i \alpha_i T_i$ where $\alpha_i \in \mathbb{R}$ are the weights of Pauli strings and $T_i \in \{I, X, Y, Z\}^{\otimes n}$ for n qubits (with the JW mapping, this is the number of spin orbitals). The role of the QPU is to compute expectation values of this Hamiltonian. In particular, for a given set of parameters $\vec{\theta}$, the QPU calculates $\left\{\langle T_i \rangle_{\vec{\theta}}\right\}$ which may then be weighted on the classical computer according to the coefficients α_i to form $\langle H \rangle_{\vec{\theta}}$. Work has been done to reduce the number of measurements at this step based on the properties of $\{T_i\}$ [41]. Additionally, Ref. [57] developed techniques that use symmetry to reduce the number of qubits for certain systems including H_2, however the transformations used are not compatible with our framework.

Despite substantial progress, accurately performing VQEs on current and near-term hardware is still challenging for a number of reasons. Firstly, calls to the objective function $f(\vec{\theta})$ are inherently probabilistic, meaning one must use an optimization algorithm to minimize $f(\vec{\theta})$ that is resistant to noise. Additionally, the number of variational parameters used to construct the variational form can potentially limit the performance of an optimization algorithm in finding a suitable minimum for the estimated ground state energy. Furthermore, the depth of circuits and number of CNOT gates are costly resources when executing circuits on a real QPU, so creating ansätze that efficiently use these resources is paramount.

B. Symmetry Preserving Ansätze

The minimal number of parameters for a general Hilbert space of n qubits (i.e. n spin orbitals) is given by

$2(2^n-1)$. However if symmetries are present, they can be used to reduce this parameter count, sometimes dramatically. For example, if particle number is conserved, then the number of parameters reduces to $2(\binom{n}{m}-1)$, where $m < n$ is the number of fermions. If the Hamiltonian also respects time-reversal symmetry, as is often the case in molecular problems, then the energy eigenstates can be chosen to have only real coefficients, which further reduces the parameter count by a factor of 2. Total spin s and spin projection s_z are also commonly conserved, in which case the parameter count becomes [22]

$$\sum_{k=0}^{m/2-s} \binom{n/2}{k}\binom{n/2-k}{m-2k} \times \frac{(2s+1)(m-2k)!}{(m/2-k-s)!(m/2-k+s+1)!}. \tag{2}$$

Choosing a particular symmetry-preserving ansatz amounts to restricting the search space of the optimizer to a smaller region of the Hilbert space. This in turn reduces the number of variational parameters, lessening the burden on the classical optimizer. We show below that it also lessens the demands on the QPU by cutting down the number of function calls in the optimization process and reducing the circuit depth.

Using a symmetry-preserving ansatz contrasts sharply with using ansätze that rely on ad hoc arguments [5, 58]. These ansätze usually attempt to express all possible states in the Hilbert space, including the ground state of the Hamiltonian in question. The conventional wisdom here is that more variational parameters result in more flexibility in producing the target states. However, too many variational parameters can result in large circuit depths and challenging conditions for the optimizer. It may also not be possible to exactly represent the ground state with these types of ansätze if they are not designed to uniformly cover the relevant part of the Hilbert space.

In this work, we primarily focus on the "ASWAP" ansatz, which is constructed using gates that preserve the number of excitations in a state. In the case of a molecular Hamiltonian, this corresponds to fixing the number of electrons occupying some number of orbitals. The gates we consider [21] are of the form

$$A(\theta,\phi) = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & \cos\theta & e^{i\phi}\sin\theta & 0 \\ 0 & e^{-i\phi}\sin\theta & -\cos\theta & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}. \tag{3}$$

Ref. [22] showed how to construct circuits that preserve particle number and spin projection S_z by tiling these A gates in a regular pattern like the one shown in Fig. 1. Preserving any value of $\langle S_z\rangle$ is accomplished by dividing the register of qubits into two halves (the upper half represents spin-up orbitals, and the lower half spin-down). The parameters for the A gates that join the two spin subspaces are then set to 0. The desired S_z subspace can be selected by changing which subset of qubits are acted on by the initial X gates. In this work, we primarily use the ASWAP ansatz that fixes $\langle S_z\rangle = 0$, except when considering excited states in Sec. VI.

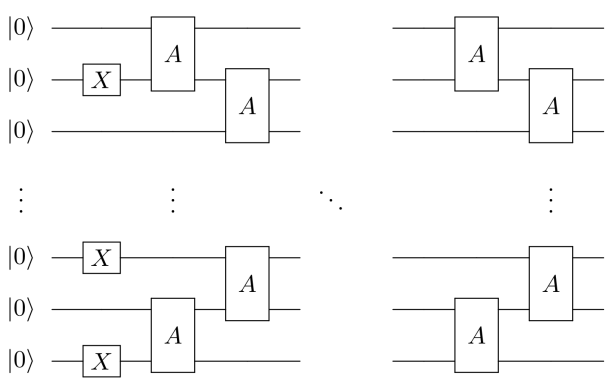

FIG. 1. Circuit used to produce the "ASWAP" ansatz. Each A gate has two separate parameters, θ and ϕ. Each A gate preserves the particle number of the state it acts on, so the entirety of the ansatz does too. The initial X gates are used to place the register in the desired particle number subspace.

The advantages of the symmetry-preserving ansätze are thus twofold. Firstly, restricting the space of states prepared by the ansatz always produces states that are consistent with physical constraints imposed by the Hamiltonian. Secondly, both the classical and quantum resources required to run the VQE are reduced. The latter is especially important when considering near-term devices.

III. PERFORMANCE COMPARISONS OF VARIOUS ANSÄTZE WITHOUT NOISE

Before testing the performance of the symmetry preserving ansätze in the presence of noise, we first demonstrate that the correct ground state can be prepared in a noiseless context. To do this, we run different instances of the VQE algorithm for H_2 under the JW mapping on four qubits using a noiseless simulator. The noiseless simulator used is the "state vector simulator" implemented in Qiskit. To perform the optimization of the variational parameters in the VQE, we use the limited memory BFGS (Broyden-Fletcher-Goldfarb-Shannon) optimization algorithm. For noiseless simulations, the number of CNOT gates in an ansatz will not affect the quality of the results. Thus, if a given ansatz is able to reach chemical accuracy, the relevant benchmark is how many function evaluations are required to do so.

Since our ansätze are symmetry preserving, the reference state ultimately determines the targeted values of $\langle N\rangle$ and $\langle S_z\rangle$. Therefore, in the case of H_2, we expect a singlet ground state, so we have $\langle N\rangle = 2$ and $\langle S_z\rangle = 0$. The ansätze that we compare against are standard ad hoc ansätze available in Qiskit [52]. Generally,

the structure of these ansätze involves interleaving layers of single-qubit rotations with entangling operations. In the case of the RY(RYRZ) ansätze, layers of single-qubit $R_Y(\theta)$ ($R_Y(\theta)R_Z(\phi)$) operations on all qubits are interleaved with CZ gates. However, since CZ and CNOT are locally equivalent, we instead count the number of CNOT operations for comparison. In all cases when executing circuits, we use the highest level of optimization available in Qiskit to reduce gate counts (i.e. transpilation). The number of such layers in the ansatz is referred to as the depth. The SwapRZ ansatz interleaves $R_Z(\theta)$ operations with parameterized SWAP operations, meaning that this ansatz preserves particle number symmetry. In contrast with our ASWAP ansatz, the SwapRZ ansatz does not in general span the full symmetry subspace, and so does not necessarily capture the ground state exactly. Additionally, the ASWAP ansatz generally has fewer parameters and CNOTs. In these ad hoc ansätze, the entangling operations performed at each layer can be chosen for the appropriate problem. For the ansätze considered here, we perform entangling operations between all pairs of qubits unless specified otherwise.

In Fig. 2, we illustrate the performance of several ad hoc ansätze, along with our ASWAP ansatz. All the ansätze considered are able to find the correct ground state energy to well below chemical accuracy (~ 1.5 milli-Hartree). One distinction between our ansatz and the ad hoc ansätze, however, is that the ASWAP ansatz requires at least five times fewer function evaluations to converge compared to any of the other considered ansätze, while simultaneously having lower error rates. For the example shown in Fig. 2, ASWAP requires 414 function calls, while RY uses 2142. This improvement is expected since the number of variational parameters is smaller than the other considered ansätze. This illustrates the key point that enforcing known symmetries of the simulated Hamiltonian can significantly reduce the computational load on the QPU, namely the number of calls to the objective function.

The metrics that we use to evaluate the resource "efficiency" of a particular ansatz are the number of variational parameters and the number of CNOT operations it contains. We summarize these resource requirements for ASWAP and for several of the ad hoc ansätze described above in Table I for 4 qubits. Generally, as expected, the number of variational parameters in the ASWAP ansätze can be small, despite being able to quickly and accurately prepare the correct ground state.

Ansätze with more parameters require more optimization steps and hence calls to the objective function. Ansätze with more CNOT gates suffer from higher error rates and are more limited by decoherence. Here, the ASWAP ansätze enforce time-reversal symmetry. The combinations of symmetries for ASWAP are not exhaustive, as more are possible. Some ansätze are parameterized by the "depth" $d \geq 1$, which is the number of times the layer is repeated. Ansätze for which a symmetry is "None" means that the ansatz does not enforce this symmetry. The RY(RZ) ansätze apply CZ operations between nearest-neighbors at each layer.

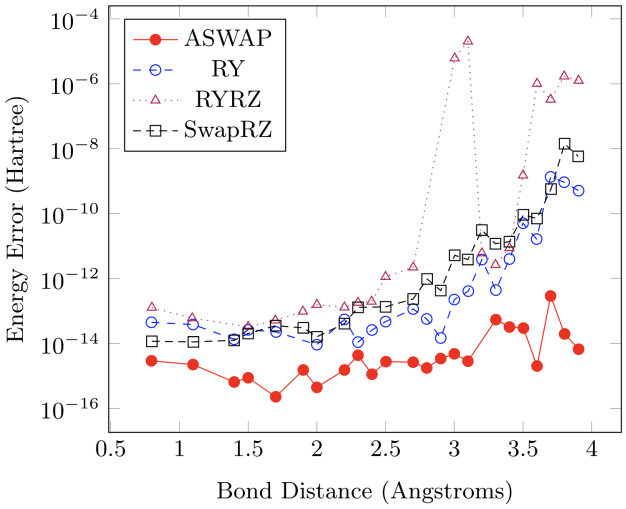

FIG. 2. Results for H_2 in a noiseless simulator with the limited memory BFGS algorithm. All results are well below chemical accuracy and sufficiently converged. To account for bad initial variational parameters, each result was obtained by running additional optimizations with randomized initial variational parameters. The lowest energy found with this method was then chosen as the final result. The median number of iterations of the BFGS optimizer for each ansatz were: ASWAP 414, RY 2142, SwapRZ 3955, RYRZ 4851.

This comparison of resources illustrates the potential of enforcing symmetries at the level of the ansatz. Not only is the final prepared state guaranteed to have certain desired properties, but enforcing this at the level of the ansatz allows for a reduced number of parameters as well as CNOT gates, compared to a more generic ansatz.

TABLE I: Comparison of ansatz resources on 4 qubits.

	N	S_z	# Pars	# CNOTs
ASWAP	0, 4	None	0	0
	1, 3	None	3	9
	2	None	5	16
	2	0	3	3
RY(d)	None	None	$4(d+1)$	$3d$
RYRZ(d)	None	None	$8(d+1)$	$3d$
SwapRZ(d)	N	None	$10d+4$	$24d$

IV. EFFECTS OF NOISE ON PERFORMANCE OF VQE WITH VARIOUS ANSÄTZE

A. Hardware Noise and Noise Model Simulation

All of the ansätze used in this paper are applicable to any general-purpose quantum processor. Nevertheless, for the purposes of noisy simulation, we will consider a noise model that is derived from information about the

IBM Q processors. In particular, we include information from the devices about measurement errors, single- and two-qubit gate errors, depolarization, and thermal relaxation errors that are dependent on individual gate times.

To establish a vocabulary, the different simulators in Qiskit that we use here are as follows. We refer to the noiseless simulator that keeps track of the state vector of the register under unitary operations as the "state vector simulator". We refer to the noisy simulator that includes the above information as the "QASM simulator" (quantum assembly language) [52]. When using the QASM simulator, we will use noise models constructed from information about the Vigo, Boeblingen, Ourense, and Johannesburg IBM Q devices.

B. Error Mitigation Techniques

When running a VQE in the presence of noise, we use three error mitigation techniques. The first of these addresses state preparation and measurement (SPAM) errors by computing and inverting a matrix T of dimension $2^n \times 2^n$ whose entries represent the probability of preparing one state and immediately measuring another. Inverting this matrix allows one to perform error mitigation on population counts of a particular experiment, thereby accounting for SPAM errors. This and similar techniques have been previously explored [59–64]. Though there is not yet a fully scalable solution for mitigating SPAM errors, each of these techniques have applicability in different contexts for near-term devices. Here, we use the implementation in Qiskit.

The second technique, known as Richardson extrapolation [29, 65, 66], involves systematically increasing the amount of error in a given computation in a controlled way so that one can then extrapolate an objective function to a point at which those errors would be reduced. In this case, the extrapolation occurs by taking each CNOT gate in a particular circuit and inserting an additional even number of CNOTs immediately after it. For instance, each CNOT gate in a given circuit would then become a sequence of an odd number of CNOT gates. In this way the logic of the circuit is unaffected, but in the presence of noise and faulty gates, more redundant CNOT gates will increase the error in the results. In Fig. 3, this procedure is shown for a variety of bond distances. In this case, we find that the extrapolated minimum energy obtained by the VQE is more accurate than the original results by as much as an order of magnitude in the energy error. Moreover, the extrapolated results are within a standard deviation of chemical accuracy for most of the interatomic distances.

The third error mitigation we use is partial symmetry enforcement during measurement. This noise mitigation relies on the fact that the true ground state has a known, definite particle number and therefore the measurement results should also have this symmetry, even in the presence of noise. This mitigation is only partially possible since the measurement process itself requires post-rotation gates, which do not themselves conserve particle number. However, for the chemical Hamiltonians we are interested in, we can always group together Pauli strings which do preserve particle number. For instance, the H_2 Hamiltonian, after the JW mapping onto 4 qubits, can be divided into 5 sets of commuting Pauli strings; one of these sets is made up of only Pauli strings that commute with the total particle number operator (terms like $IIIZ, ZZII, IZII$ etc.). During the measurement of this set, we can enforce that the particle number must match that of the ground state, discarding results which do not respect this symmetry. We also note that this procedure is not limited to the JW mapping. Under the BK mapping, for instance, the $a_i^\dagger a_i$ and $a_i^\dagger a_j^\dagger a_j a_i$ terms in a molecular Hamiltonian map to operators that are diagonal in the Pauli basis [67], as does the total particle number operator. Hence, this procedure extends to other mappings, however we cannot presently compare their performance since the ASWAP ansatz currently only applies to the JW mapping.

This process simply post-selects data which we know to positively contribute to our goal of finding the ground state. The measurements of Pauli strings that do not commute with the total particle number operator are not modified. This technique may be more applicable when the state produced on the QPU during the VQE produces expectation values that are close to the total particle number, which is the case with the ASWAP ansätze. Nevertheless, this may be applied to any ansatz as we later demonstrate. One potential drawback of this approach is the overhead introduced in the number of shots needed to accurately take expectation values. In discarding shots, it may be necessary to perform additional measurements to compensate for the reduced shots after post-selection. In this case, as we later discuss, we are limited in the number of shots by the IBM Q devices.

We employ all of these strategies in various combinations in a noisy H_2 simulation using both the RY and ASWAP ansätze. The results are shown in Fig. 4. We find that the largest improvement in performance for the RY and ASWAP ansätze come from partial symmetry enforcement and SPAM strategies, respectively. In the case of RY, the large improvement under partial symmetry enforcement is likely due to the fact that RY can prepare states that do not obey the desired symmetry, but partial symmetry enforcement corrects this by classical post-processing. In the case of ASWAP, the large improvement from SPAM is likely due to the structure of the ansatz. Specifically, the ASWAP ansatz uses logic that assumes a fixed initial state of the register. Moreover, measurement errors can in principle violate the expected value of the particle number. Both of these effects can be suppressed by SPAM error mitigation.

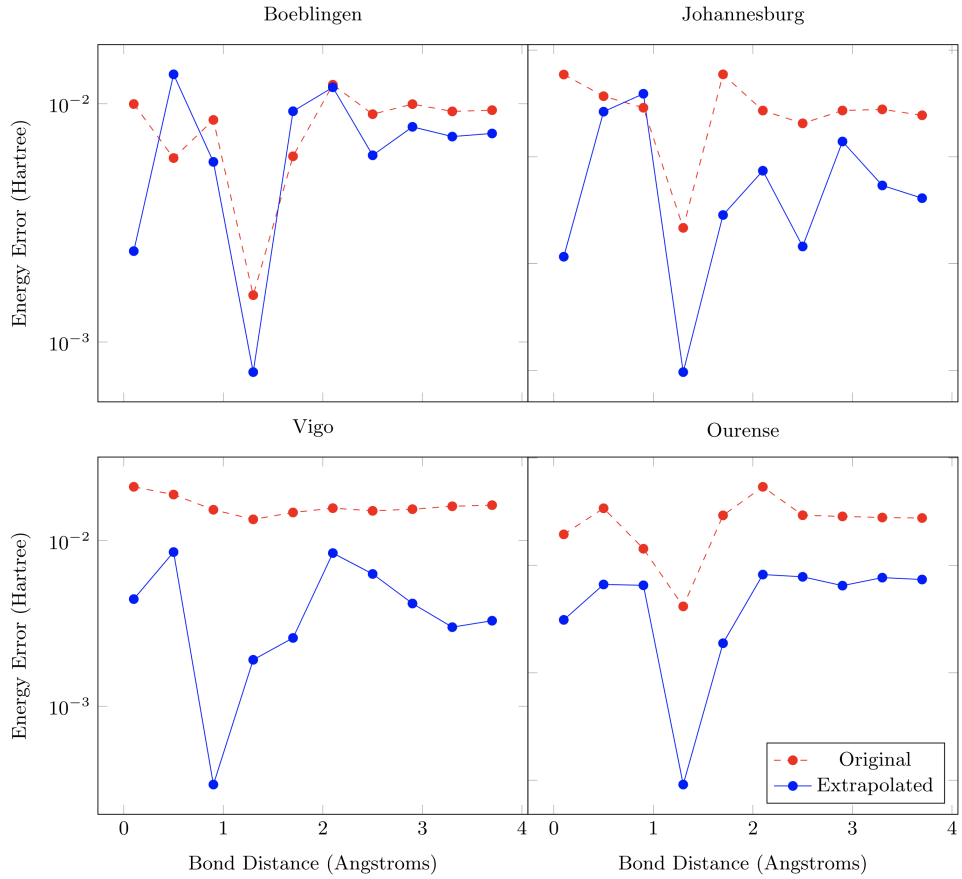

FIG. 3. Comparison of results with and without Richardson extrapolation for H_2 at varying interatomic distances for the ASWAP ansatz. Some points on the extrapolated curve are within chemical accuracy ($\sim$ 1.5 milliHartree). Richardson extrapolation provides as much two orders of magnitude improvement over a standard VQE run.

C. Noisy Optimization

Minimization of the objective function $f(\vec{\theta})$ requires an optimization algorithm that is resistant to noise. In previous works, variational quantum algorithms have been implemented using the SPSA (simultaneous perturbation stochastic approximation) [68], Adam (adaptive moment estimation) [69], COBYLA (constrained optimization by linear approximation) [70], and DIRECT (dividing rectangles) [71] algorithms. For this work, we choose the DIRECT and COBYLA algorithms for three primary reasons. First, the DIRECT algorithm has been demonstrated to be resistant to substantial amounts of noise in variational quantum contexts [31]. Second, DIRECT is a global optimization algorithm that rarely gets trapped in local minima. In this sense, DIRECT sidesteps the need to consider carefully chosen (and potentially repeated) initial conditions for the variational parameters. Finally, though DIRECT does not converge quickly, it gets within the vicinity of the correct solution quickly. This trade-off is beneficial for the case of a noisy objective function since high levels of precision in the objective function are limited by statistical and hardware errors. We also choose the COBYLA optimization algorithm for similar reasons, namely its ability to find an accurate estimate of the ground state in a short number of function calls and resilience to noise. In previous experiments, it has been shown that the number of shots limits the resulting accuracy of the final energy estimate [5]. For this reason, we performed the maximum number of shots allowed by IBM devices (8192).

For our noisy simulations, the backend chosen is the QASM simulator, which uses a noise model constructed to mimic a variety of IBM Q processors. We use the DIRECT and COBYLA optimization algorithms. In all simulations using DIRECT, we limit the optimization to a budget of 100 or 300 calls to the objective function. For the COBYLA optimizer, on average the optimizer converges with about 400 calls to the objective function.

D. Results for noisy VQE simulations

In Fig. 5 we show the results of running the VQE for a variety of bond distances of H_2 for the best-performing symmetry preserving and ad hoc ansätze. We find that

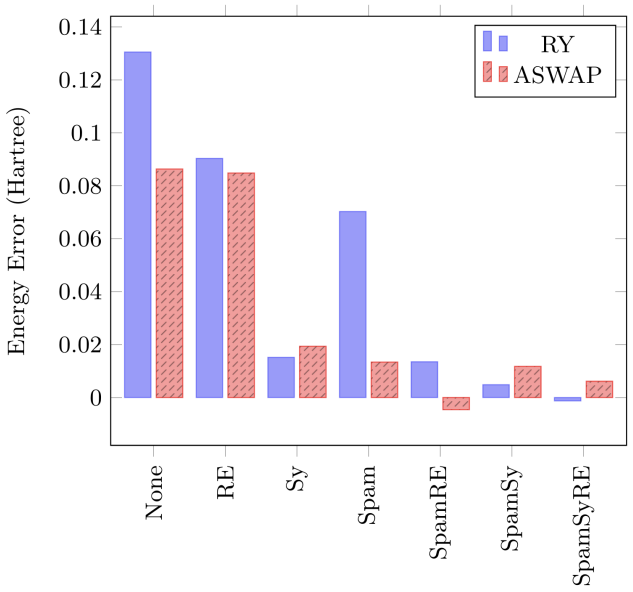

FIG. 4. Performance of two competing ansätze with different combinations of error mitigation strategies for H_2 at equilibrium. Each of these strategies is introduced in Sec. IV B. The abbreviations for the strategies are RE (Richardson Extrapolation), Sy (Partial Symmetry Enforcement), Spam (SPAM), or combinations of any of these. We note here that the ASWAP ansatz performs better even in the absence of any error mitigation strategies. This underscores the fact that the reduction in quantum and classical resources due to imposing symmetries at the ansatz level leads to improvements in the energy error. Additionally, some values of the energy error are below zero, which violates the variational principle. This is due to our implementation of Richardson Extrapolation, which can potentially undershoot the value of the ground state energy. This is also potentially due to shot noise, since our simulations rely on a finite number of samples from the QPU.

for smaller bond distances, the ASWAP ansatz performs better in terms of finding the correct ground state energy than the ad hoc ansätze. At larger bond distances, it appears that the ad hoc ansätze perform better; however, as we show in the next section, this is actually not the case. The reason is because in this regime, the ad hoc ansätze are actually converging to an excited state that is nearly degenerate with the true ground state. All the ansätze have greater difficulty in identifying the true ground state at large bond distance because correlations are stronger here (the true ground state is a singlet). On the other hand, the excited state found by the ad hoc ansätze is a (triplet) product state, which is easy for the ad hoc algorithms to identify. Because ASWAP has the correct symmetry quantum numbers built in, it does not misidentify the ground state, and the error at large bond distance is larger because of the correlations in this state.

E. Experimental Simulation of H_2

To more rigorously demonstrate the effectiveness of these techniques in the presence of noise, we use the physical IBM superconducting qubits to validate the results. In Fig. 7 we show the dissociation of H_2 using several different devices. In each case, the resulting energy shows good agreement with the exact result, and the shape of the dissociation curve is generally preserved.

V. EFFECT OF NOISE ON SYMMETRY PRESERVATION

In addition to reducing computational resources and error levels, the ASWAP ansatz also correctly preserves the desired symmetry throughout the entire dissociation curve. This is illustrated in Fig. 6 for the case of H_2. While all ansätze considered remain reasonably close to the correct particle number, $\langle N \rangle = 2$, only ASWAP succeeds in finding the right value of $\langle S_z \rangle$ at all bond distances. In contrast, the RY ansätze find another nearly degenerate solution that has an incorrect value of S_z. This is a very important advantage of our ansätze when considering quantum chemistry problems, because with other ansätze there is no way to control which solution is found when there are approximate degeneracies aside from adding penalty terms to the objective function to find different symmetry states. Even with that approach, it is not clear how one would guarantee consistent eigenvalues for the symmetries throughout the whole curve with another ansatz. The ASWAP ansatz allows us to specify exactly which of these solutions we want to find.

A. Total Spin Ansatz Performance

So far, this work has only focused on the ASWAP ansatz from Ref. [22]. That work also presented another ansatz, referred to as the E_n ansätze, where n is the number of qubits. Unlike any of the other ansätze in the present work, it has the ability to preserve the total spin of the trial state. The E_n ansatz is constructed by first writing out a general state vector in the appropriate total spin subspace, with the coefficients parameterized by hyperspherical coordinates. Then a unitary that produces these general states of fixed total spin starting from $|0\rangle^{\otimes n}$ is constructed and decomposed in terms of Toffoli gates using Gray codes [72]. The resulting ansätze can then be used to prepare arbitrary states with the appropriate spin projection, total spin, and particle number symmetries. The details of this construction can be found in Ref. [22].

The E_n ansätze have larger gate depths, and they are generally more challenging to implement on noisy hardware. In Fig. 9 we compare the performance of the E_4 and ASWAP ansätze for simulating H_2 at equilibrium. To accomplish this, we take the noise model from the

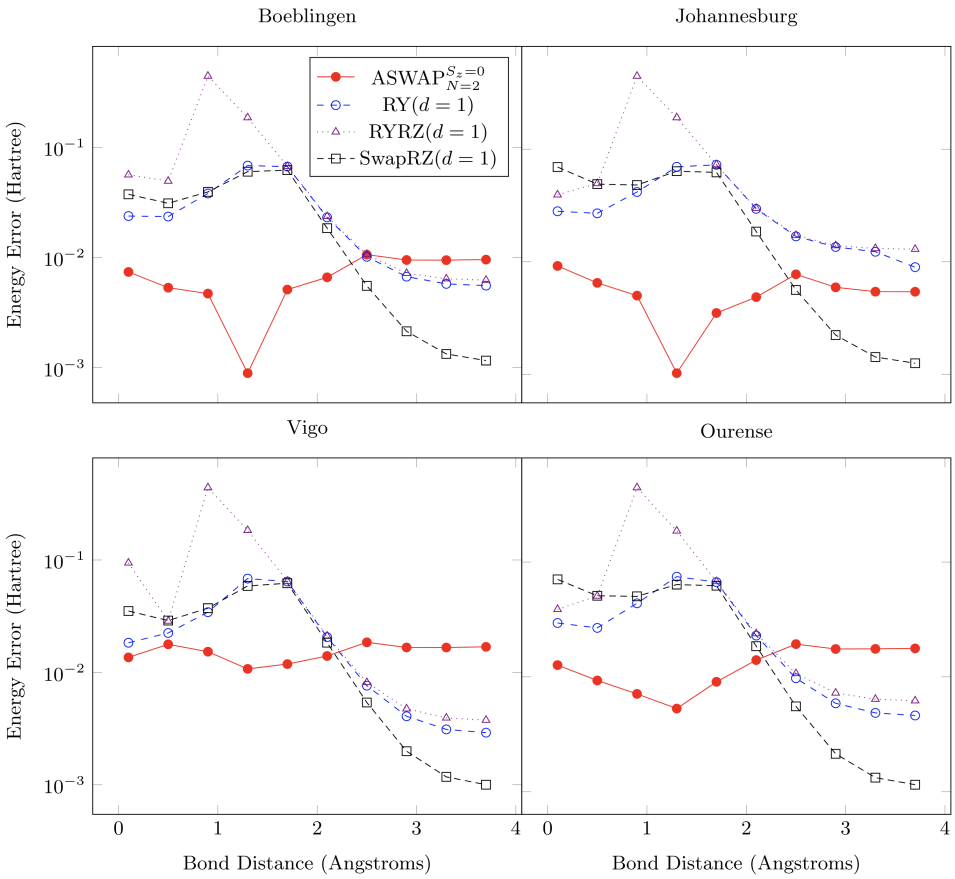

FIG. 5. Comparison of the symmetry-preserving ASWAP ansatz with several of the best competing ad hoc ansätze, for a variety of noise models corresponding to real devices. In all cases, the ASWAP ansatz performs better at smaller bond distances. This is because for large bond distances, the ad hoc ansätze find a degenerate, separable ground state, whereas the ASWAP ansatz consistently finds the non-trivial ground state for all bond distances. All results in this figure have SPAM error mitigation applied.

IBMQ Vigo device and uniformly stretch the T_1 and T_2 times for all of the qubits, so as to simulate a device with less noise. T_1 is the energy relaxation time, associated with decay from the $|1\rangle$ to the $|0\rangle$ state. T_2 is the decoherence time, associated with the loss of the phase of the quantum state. From these simulations, we confirm that the performance of the E_4 gate is limited by the $T_{1/2}$ times. Moreover, we find that these times would need to be ~ 4 times longer than they currently are to perform as well as the ASWAP ansatz with current $T_{1/2}$ times. Hence, we expect that the E_4 ansatz will be useful for preserving total spin, but is not yet currently viable. Nevertheless, there may be more complex molecules for which the ability to preserve S^2 will have a greater impact on the results and offset the challenges of the larger circuit depth.

B. Application to Heisenberg Model

So far we have only focused on molecular systems, but now we will turn our attention to the Antiferromagnetic Heisenberg model (with $J > 0$). The Hamiltonian is given by

$$H = J \sum_{\{i,j\}\in E(L)} (X_i X_j + Y_i Y_j + Z_i Z_j) + B \sum_{i\in V(L)} Z_i \tag{4}$$

where J is the interaction strength, B is the magnetic field strength, and $E(L)$, $V(L)$ are the edge and vertex sets of the lattice L, respectively. Here, we choose the 2×3 lattice. For this system, the ground state has a well-defined particle number which varies with different values of J/B. Here the particle number is defined as $\sum_{i\in V(L)}(1 - \langle Z_i\rangle)/2$. We have chosen this definition of the particle number since (in the computational basis) its diagonal entries correspond to the Hamming weight of the corresponding basis state. This is more amenable to the application of ASWAP operations. In Fig. 8 we

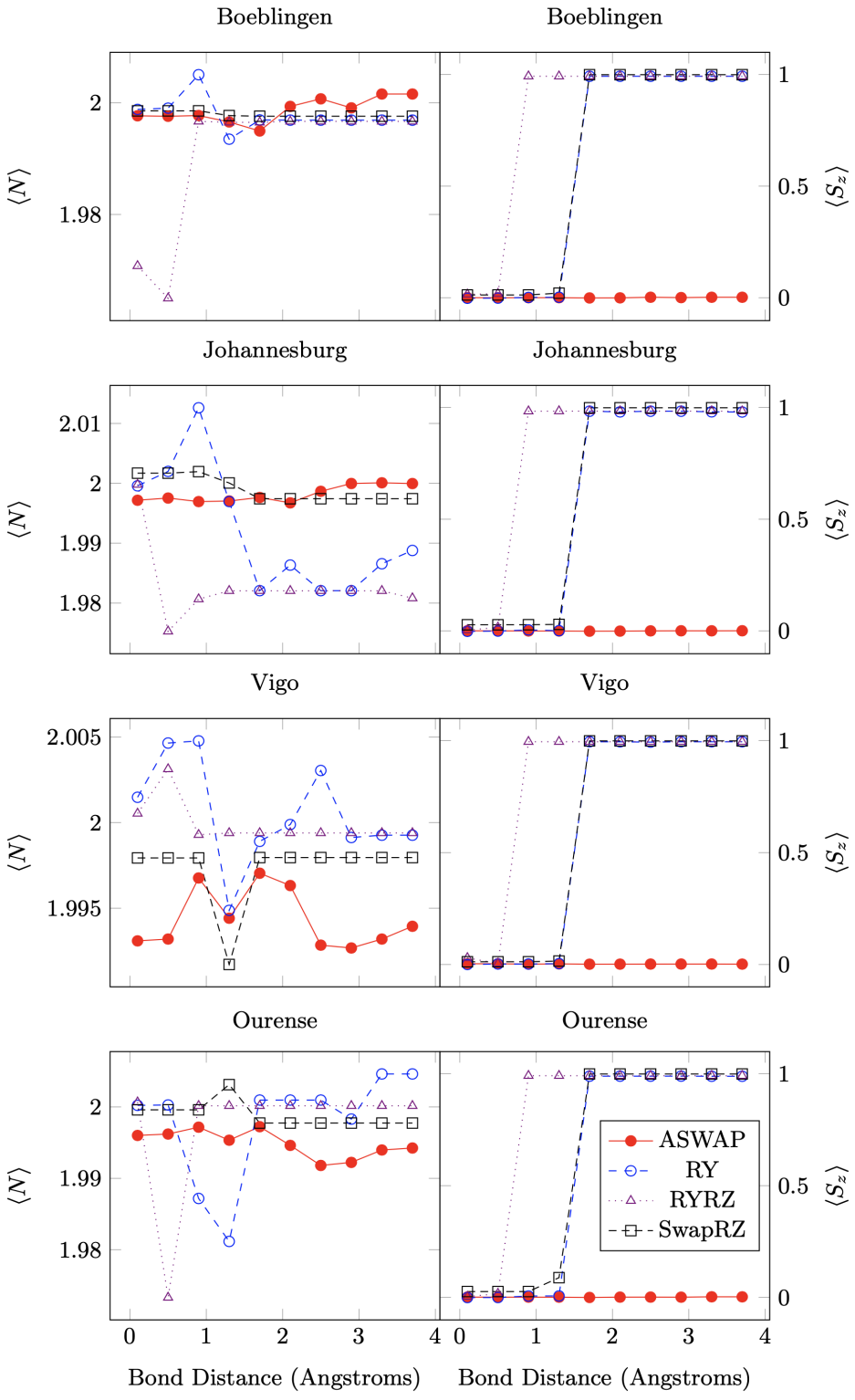

FIG. 6. Comparisons of preserved symmetries ($\langle N \rangle, \langle S_z \rangle$) in ASWAP, RY, and RYRZ ansätze. The ASWAP ansatz consistently preserves each desired symmetry throughout the dissociation, and any deviation, in this case, is due to the noise model. On the other hand, deviations in the RY and RYRZ cases can also be attributed to the inability of the ansatz to target a particular subspace. This is most prominently shown in the case of the $\langle S_z \rangle$ operator (right column). For the RY ansätze, the change in the resulting expectation values around 1 Angstrom is due to the algorithm finding another, nearly degenerate solution for the ground state that has an incorrect value of $\langle S_z \rangle$. The fact that the ad hoc ansätze produce values of $\langle S_z \rangle = +1$ and not -1 is dependent on both the optimization algorithm used, as well as the depth and connectivity used in the ansatz. Results using the DIRECT algorithm (shown here) consistently produce $\langle S_z \rangle = +1$ whereas COBYLA occasionally finds solutions with $\langle S_z \rangle = -1$. When using COBYLA, the value of $\langle S_z \rangle$ depends on the depth and connectivity of the ansatz used. This underscores the importance of using an ansatz that preserves known symmetries so that one can guarantee consistent results. All results in this figure have SPAM error mitigation applied.

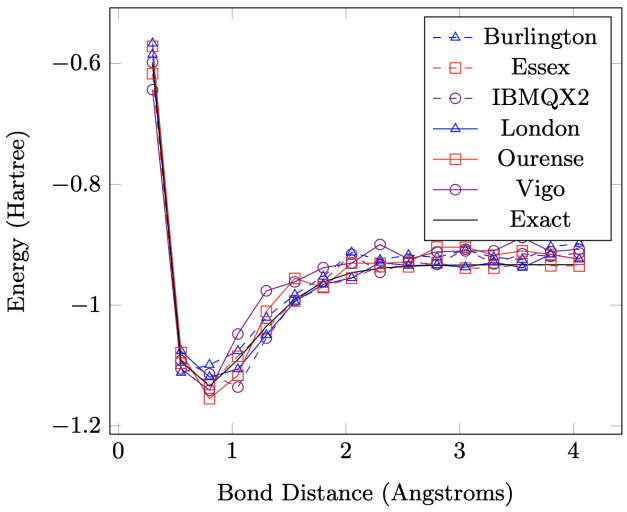

FIG. 7. Comparison of experimental VQE results for the dissociation of H_2 using IBM Q devices using the ASWAP ansatz with $S_z = 0$, $N = 2$, and time-reversal symmetry enforced. In each case, the shape of the results from hardware match the general shape of the exact result, even at longer bond distances. All points (except the exact results) have SPAM and Richardson Extrapolation error mitigation applied. This highlights the robustness of the ASWAP ansatz to a variety of noise levels in an experimental setting.

show that again the ASWAP ansatz is able to successfully find the ground state with the correct particle number, whereas the RY ansatz produces results that deviate from the correct particle number. Moreover, in some regions, the ASWAP ansatz produces better results in terms of the energy error. These results show that if the symmetry of the ground state is known a priori, then enforcing it at the level of the ansatz can lead to a reduction in resource requirements and error rates. It is also important to note that even if the symmetry of the ground state is not known beforehand, one can still try multiple ansatz symmetries in independent VQEs run on different QPUs in parallel. It should also be noted that the ad hoc ansätze can explore different symmetries during a single VQE implementation run on one QPU. This potentially makes such ansätze appealing if nothing about the symmetries is known, and access to multiple QPUs is limited.

VI. EXCITED STATES

An important aspect of our symmetry-preserving ansätze is the ability to find excited states by targeting subspaces of the Hilbert space that have the appropriate symmetry, provided this symmetry differs from that of the ground state. Other methods of calculating excited states using VQEs exist, but they require either iteratively running multiple VQEs to determine excited states [32] or executing additional circuits [6, 33]. Nevertheless, each of these approaches has its own advantages. For example, Ref. [33] does not require multiple optimizations and is robust to noise. Hence, in terms of optimization, no additional effort is required to obtain excited states.

A notable feature of our method is that the search for excited states does not depend on having already obtained the ground state. In fact, we could compute both ground and excited states simultaneously by running separate VQEs in parallel, or skip the ground state if we are only interested in excited states. Since our method amounts to simply changing the ansätze for different excited states, finding an excited state is as easy as finding the ground state. One possible disadvantage of our approach is that depending on the system chosen, it may not be the case that one can uniquely specify all excited-states by changing symmetries alone. Therefore, the optimal approach may be a combination of our techniques and some of the previously introduced ones.

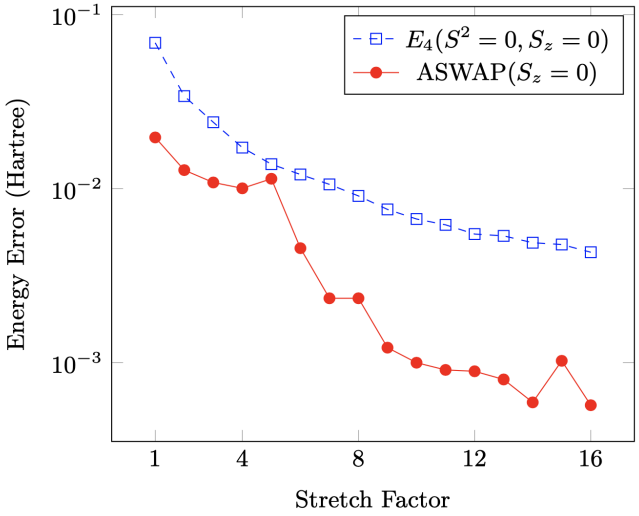

FIG. 9. Performance of the E_4 and ASWAP ansätze for H_2 at equilibrium for a range of stretch factors. The stretch factors here are the ratios of the original and simulated T_1 and T_2 times for the noise model. A stretch factor of 1 corresponds to the original noise model. Here we chose the noise model for the IBMQ Vigo device. We find that improving the T_1 and T_2 times of the device by a factor of roughly 4 makes the E_4 ansatz perform as well as the ASWAP ansatz does today. Since the only error mitigation technique used here is SPAM error mitigation, we find that roughly an order of magnitude improvement of the T_1 and T_2 times will allow the ASWAP ansatz to reach chemical accuracy without Richardson extrapolation.

A demonstration of excited state calculations using our ASWAP ansatz is shown for the H_2 molecule for several excited states in Fig. 10. Note that some of the exact results do not have corresponding simulation results. This is because our method does not, in general, allow one to produce all of the higher excited states, only those directly accessible by partitioning the Hilbert space according to symmetries. In all cases considered, we are able to

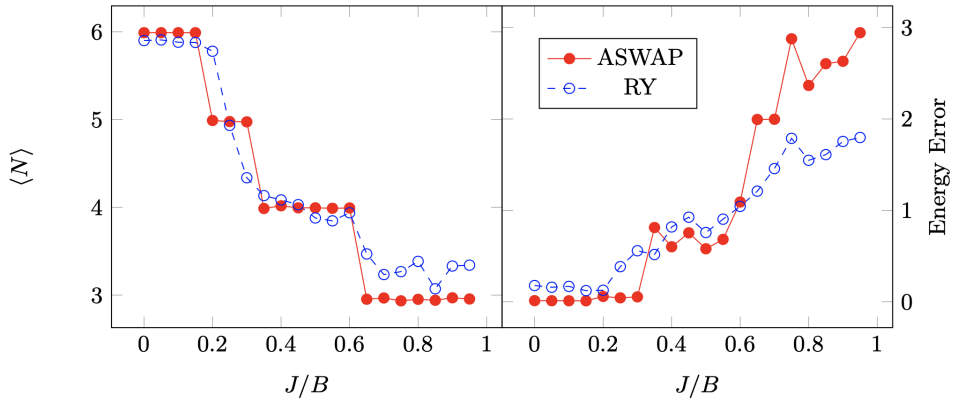

FIG. 8. Comparison of VQE results using ASWAP and RY ansätze for the Heisenberg model on a 2×3 lattice. Since the particle number cannot be fixed to a constant value with the RY ansatz, it only sometimes produces the correct result. On the other hand, by fixing the particle number in each case to the correct value, the ASWAP ansatz is able to obtain better results. To improve the RY results, we use 3 layers of entanglers and rotations. Here, we use the IBMQ Vigo noise model and include SPAM error mitigation.

prepare 5 excited states (in addition to the ground state) out of the possible 10 (not counting degeneracies). The error rates here are sufficiently low so that one may correctly order the eigenstates according to their energies.

VII. OUTLOOK

The key development of Ref. [22] was that the ASWAP tiling procedure produces ansätze that can be used to prepare any state in the given symmetry subspace. However, the tiling procedure used has the requirement of using nearest-neighbor connectivity. This raises the possibility of other connectivities having this same property, namely that they span symmetry subspaces. Though beyond the scope of the current work, future implementations of ASWAP ansätze could be hardware tailored by tiling A gates according to the connectivity of the device. Such a result would alleviate the need to insert additional SWAP operations to account for non-nearest neighbor operations. In this case, one would need to check that the ansatz is able to capture all states in the symmetry subspace. The resulting ansatz could be considered a hybrid between hardware efficient (taking into account the device connectivity) and symmetry preserving (only utilizing A gates) ansätze. Such considerations may become more advantageous as device architectures become more complex.

VIII. CONCLUSIONS

In this work, we have demonstrated that the deleterious effects of noise on a VQE can be mitigated by exploiting the symmetries of the Hamiltonian being simulated. We achieve this by performing simulations of the noiseless and noisy systems for a variety of ansätze, utilizing the most current error mitigation techniques. We also test our methods on real IBM devices. The results of these simulations and experiments indicate that, with or without noise, using state preparation circuits that preserve the relevant symmetries of the problem reduces computational resources and error rates compared to more ad hoc approaches. In addition, we have shown that the built-in symmetry preservation of our ansätze allows us to find excited states without having to first find the ground state and without using longer state preparation circuits.

ACKNOWLEDGEMENTS

S. E. E. acknowledges support from the US Department of Energy (Award No. DE-SC0019318). E. B. and N. J. M. acknowledge support from the US Department of Energy (Award No. DE-SC0019199). 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. Oak Ridge National Laboratory manages access to the IBM Q System as part of the IBM Q Network.

Appendix A OPERATORS WITHOUT SPAM MITIGATION

In Fig. 11 we show the same information as in Fig. 6, except without using SPAM error mitigation. In terms of preserving symmetries, SPAM error mitigation helps to reduce the errors in estimating $\langle N \rangle$ that are due to state preparation and measurement. Generally, we see that without SPAM error mitigation $\langle N \rangle$ is underestimated. This is likely due to T_1 relaxation during state preparation and measurement.

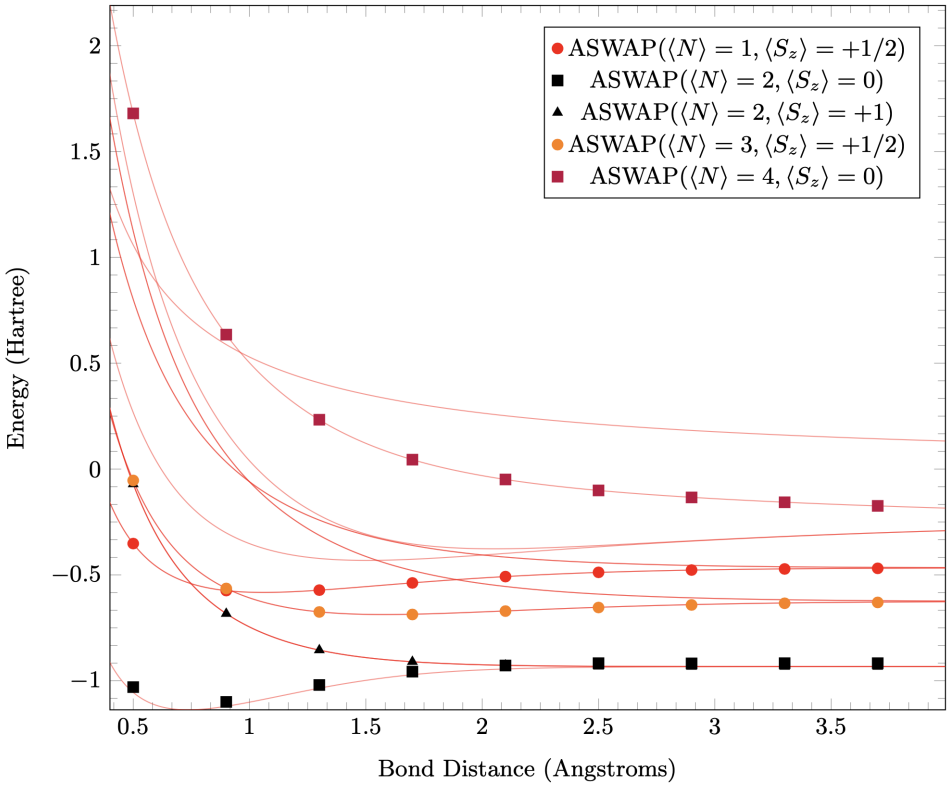

FIG. 10. The dissociation curve for a variety of excited states for H_2 calculated using the ASWAP ansatz. Dotted points correspond to results obtained using simulations of the Vigo IBMQ quantum processor, where different colors correspond to different symmetries enforced by the ASWAP ansatz (here, the particle number $\langle N\rangle$ and spin projection $\langle S_z\rangle$). Lines correspond to the different excited states as a function of the bond distance calculated using exact diagonalization, where different colors correspond to different excited states. In all cases, we find strong agreement between the results from noisy simulation and exact diagonalization. All results in this figure have SPAM error mitigation applied.

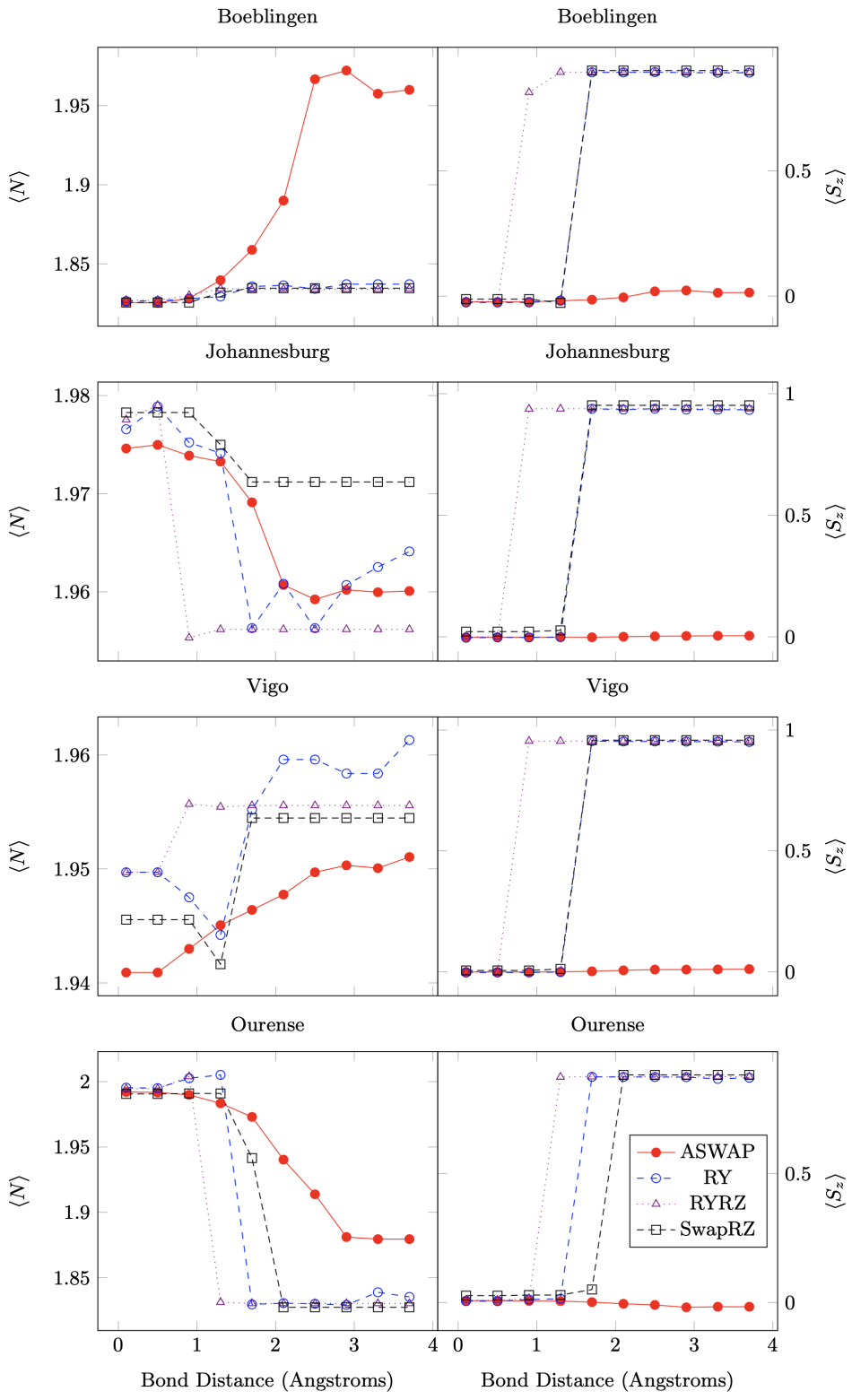

FIG. 11. Identical to Fig. 6, except with no SPAM error mitigation. Comparing with Fig. 6, SPAM error mitigation seems to consistently help preserve $\langle N \rangle$ for all ansätze.

[1] Richard P. Feynman, "Simulating physics with computers," International Journal of Theoretical Physics **21**, 467–488 (1982).

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

[3] 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, B. Campbell, Y. Chen, Z. Chen, B. Chiaro, A. Dunsworth, A. G. Fowler, E. Jeffrey, E. Lucero, A. Megrant, J. Y. Mutus, M. Neeley, C. Neill, C. Quintana, D. Sank, A. Vainsencher, J. Wenner, T. C. White, P. V. Coveney, P. J. Love, H. Neven, A. Aspuru-Guzik, and J. M. Martinis, "Scalable quantum simulation of molecular energies," Phys. Rev. X **6**, 031007 (2016).

[4] Jarrod R McClean, Jonathan Romero, Ryan Babbush, and Alán Aspuru-Guzik, "The theory of variational hybrid quantum-classical algorithms," New J. Phys **18**, 023023 (2016).

[5] Abhinav Kandala, Antonio Mezzacapo, Kristan Temme, Maika Takita, Markus Brink, Jerry M Chow, and Jay M Gambetta, "Hardware-efficient variational quantum eigensolver for small molecules and quantum magnets," Nature **549**, 242–246 (2017).

[6] 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).

[7] John Preskill, "Quantum Computing in the NISQ era and beyond," Quantum **2**, 79 (2018).

[8] Edward Farhi, Jeffrey Goldstone, and Sam Gutmann, "A quantum approximate optimization algorithm," arXiv preprint arXiv:1411.4028 (2014).

[9] G Pagano, A Bapat, P Becker, KS Collins, A De, PW Hess, HB Kaplan, A Kyprianidis, WL Tan, C Baldwin, *et al.*, "Quantum approximate optimization with a trapped-ion quantum simulator," arXiv preprint arXiv:1906.02700 (2019).

[10] Jonathan Romero and Alan Aspuru-Guzik, "Variational quantum generators: Generative adversarial quantum machine learning for continuous distributions," arXiv preprint arXiv:1901.00848 (2019).

[11] Vicente Leyton-Ortega, Alejandro Perdomo-Ortiz, and Oscar Perdomo, "Robust implementation of generative modeling with parametrized quantum circuits," arXiv preprint arXiv:1901.08047 (2019).

[12] D. Zhu, N. M. Linke, M. Benedetti, K. A. Landsman, N. H. Nguyen, C. H. Alderete, A. Perdomo-Ortiz, N. Korda, A. Garfoot, C. Brecque, L. Egan, O. Perdomo, and C. Monroe, "Training of quantum circuits on a hybrid quantum computer," Science Advances **5** (2019), 10.1126/sciadv.aaw9918.

[13] Marcello Benedetti, Delfina Garcia-Pintos, Oscar Perdomo, Vicente Leyton-Ortega, Yunseong Nam, and Alejandro Perdomo-Ortiz, "A generative modeling approach for benchmarking and training shallow quantum circuits," npj Quantum Information **5**, 1–9 (2019).

[14] Guillaume Verdon, Michael Broughton, Jarrod R McClean, Kevin J Sung, Ryan Babbush, Zhang Jiang, Hartmut Neven, and Masoud Mohseni, "Learning to learn with quantum neural networks via classical neural networks," arXiv preprint arXiv:1907.05415 (2019).

[15] Ian D. Kivlichan, Jarrod McClean, Nathan Wiebe, Craig Gidney, Alán Aspuru-Guzik, Garnet Kin-Lic Chan, and Ryan Babbush, "Quantum simulation of electronic structure with linear depth and connectivity," Phys. Rev. Lett. **120**, 110501 (2018).

[16] Tyson Jones, Suguru Endo, Sam McArdle, Xiao Yuan, and Simon C. Benjamin, "Variational quantum algorithms for discovering hamiltonian spectra," Phys. Rev. A **99**, 062304 (2019).

[17] Kosuke Mitarai, Tennin Yan, and Keisuke Fujii, "Generalization of the output of a variational quantum eigensolver by parameter interpolation with a low-depth ansatz," Phys. Rev. Applied **11**, 044087 (2019).

[18] Robert M. Parrish, Edward G. Hohenstein, Peter L. McMahon, and Todd J. Martínez, "Quantum computation of electronic transitions using a variational quantum eigensolver," Phys. Rev. Lett. **122**, 230401 (2019).

[19] Eric Anschuetz, Jonathan Olson, Alán Aspuru-Guzik, and Yudong Cao, "Variational quantum factoring," in *International Workshop on Quantum Technology and Optimization Problems* (Springer, 2019) pp. 74–85.

[20] Nikolaj Moll, Panagiotis Barkoutsos, Lev S Bishop, Jerry M Chow, Andrew Cross, Daniel J Egger, Stefan Filipp, Andreas Fuhrer, Jay M Gambetta, Marc Ganzhorn, Abhinav Kandala, Antonio Mezzacapo, Peter Mller, Walter Riess, Gian Salis, John Smolin, Ivano Tavernelli, and Kristan Temme, "Quantum optimization using variational algorithms on near-term quantum devices," Quantum Science and Technology **3**, 030503 (2018).

[21] Panagiotis Kl. Barkoutsos, Jerome F. Gonthier, Igor Sokolov, Nikolaj Moll, Gian Salis, Andreas Fuhrer, Marc Ganzhorn, Daniel J. Egger, Matthias Troyer, Antonio Mezzacapo, Stefan Filipp, and Ivano Tavernelli, "Quantum algorithms for electronic structure calculations: Particle-hole hamiltonian and optimized wavefunction expansions," Phys. Rev. A **98**, 022322 (2018).

[22] B. T. Gard, L. H. Zhu, G. S. Barron, N. J. Mayhall, S. E. Economou, and E. Barnes, "Efficient symmetry-preserving state preparation circuits for the variational quantum eigensolver algorithm," Npj Quantum Information **6** (2020), 10.1038/s41534-019-0240-1.

[23] Sukin Sim, Peter D. Johnson, and Aln Aspuru-Guzik, "Expressibility and entangling capability of parameterized quantum circuits for hybrid quantum-classical algorithms," Advanced Quantum Technologies **2**, 1900070 (2019).

[24] Panagiotis Kl Barkoutsos, Giacomo Nannicini, Anton Robert, Ivano Tavernelli, and Stefan Woerner, "Improving variational quantum optimization using cvar," arXiv preprint arXiv:1907.04769 (2019).

[25] Anton Robert, Panagiotis Kl Barkoutsos, Stefan Woerner, and Ivano Tavernelli, "Resource-efficient quantum algorithm for protein folding," arXiv preprint arXiv:1908.02163 (2019).

[26] Suguru Endo, Qi Zhao, Ying Li, Simon Benjamin, and Xiao Yuan, "Mitigating algorithmic errors in a hamilto-

nian simulation," Phys. Rev. A **99**, 012334 (2019).
[27] Yordan S. Yordanov and Crispin H. W. Barnes, "Efficient quantum circuits for quantum computational chemistry," arxiv preprint arXiv: 2005.14475 (2020).
[28] Ryan LaRose, Arkin Tikku, Étude ONeel-Judy, Lukasz Cincio, and Patrick J Coles, "Variational quantum state diagonalization," npj Quantum Information **5**, 1–10 (2019).
[29] E. F. Dumitrescu, A. J. McCaskey, G. Hagen, G. R. Jansen, T. D. Morris, T. Papenbrock, R. C. Pooser, D. J. Dean, and P. Lougovski, "Cloud quantum computing of an atomic nucleus," Phys. Rev. Lett. **120**, 210501 (2018).
[30] N. Klco, E. F. Dumitrescu, A. J. McCaskey, T. D. Morris, R. C. Pooser, M. Sanz, E. Solano, P. Lougovski, and M. J. Savage, "Quantum-classical computation of schwinger model dynamics using quantum computers," Phys. Rev. A **98**, 032331 (2018).
[31] Christian Kokail, Christine Maier, Rick van Bijnen, Tiff Brydges, Manoj K Joshi, Petar Jurcevic, Christine A Muschik, Pietro Silvi, Rainer Blatt, Christian F Roos, *et al.*, "Self-verifying variational quantum simulation of lattice models," Nature **569**, 355–360 (2019).
[32] Oscar Higgott, Daochen Wang, and Stephen Brierley, "Variational Quantum Computation of Excited States," Quantum **3**, 156 (2019).
[33] Pauline J Ollitrault, Abhinav Kandala, Chun-Fu Chen, Panagiotis Kl Barkoutsos, Antonio Mezzacapo, Marco Pistoia, Sarah Sheldon, Stefan Woerner, Jay Gambetta, and Ivano Tavernelli, "Quantum equation of motion for computing molecular excitation energies on a noisy quantum processor," arxiv preprint arXiv: 1910.12890 (2019).
[34] Kunal Sharma, Sumeet Khatri, M. Cerezo, and Patrick J. Coles, "Noise resilience of variational quantum compiling," arxiv preprint arXiv: 1908.04416 (2019).
[35] Michael Khn, Sebastian Zanker, Peter Deglmann, Michael Marthaler, and Horst Wei, "Accuracy and resource estimations for quantum chemistry on a near-term quantum computer," Journal of Chemical Theory and Computation **15**, 4764–4780 (2019), pMID: 31403781.
[36] Jarrod R. McClean, Sergio Boixo, Vadim N. Smelyanskiy, Ryan Babbush, and Hartmut Neven, "Barren plateaus in quantum neural network training landscapes," Nat. Commun. **9**, 4812 (2018).
[37] Edward Grant, Leonard Wossnig, Mateusz Ostaszewski, and Marcello Benedetti, "An initialization strategy for addressing barren plateaus in parametrized quantum circuits," Quantum **3**, 214 (2019).
[38] Robert M Parrish, Joseph T Iosue, Asier Ozaeta, and Peter L McMahon, "A jacobi diagonalization and anderson acceleration algorithm for variational quantum algorithm parameter optimization," arXiv preprint arXiv:1904.03206 (2019).
[39] Robert M Parrish, Edward G Hohenstein, Peter L McMahon, and Todd J Martinez, "Hybrid quantum/classical derivative theory: Analytical gradients and excited-state dynamics for the multistate contracted variational quantum eigensolver," arXiv preprint arXiv:1906.08728 (2019).
[40] Vladyslav Verteletskyi, Tzu-Ching Yen, and Artur F Izmaylov, "Measurement optimization in the variational quantum eigensolver using a minimum clique cover," arXiv preprint arXiv:1907.03358 (2019).
[41] Artur F. Izmaylov, Tzu-Ching Yen, Robert A. Lang, and Vladyslav Verteletskyi, "Unitary partitioning approach to the measurement problem in the variational quantum eigensolver method," Journal of Chemical Theory and Computation **16**, 190–195 (2020), pMID: 31747266.
[42] Pranav Gokhale, Olivia Angiuli, Yongshan Ding, Kaiwen Gui, Teague Tomesh, Martin Suchara, Margaret Martonosi, and Frederic T Chong, "Minimizing state preparations in variational quantum eigensolver by partitioning into commuting families," arXiv preprint arXiv:1907.13623 (2019).
[43] Max Wilson, Sam Stromswold, Filip Wudarski, Stuart Hadfield, Norm M Tubman, and Eleanor Rieffel, "Optimizing quantum heuristics with meta-learning," arXiv preprint arXiv:1908.03185 (2019).
[44] Xavier Bonet-Monroig, Ryan Babbush, and Thomas E O'Brien, "Nearly optimal measurement scheduling for partial tomography of quantum states," arXiv preprint arXiv:1908.05628 (2019).
[45] Artur F Izmaylov, Tzu-Ching Yen, and Ilya G Ryabinkin, "Revising the measurement process in the variational quantum eigensolver: is it possible to reduce the number of separately measured operators?" Chemical science **10**, 3746–3755 (2019).
[46] 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).
[47] Sam McArdle, Xiao Yuan, and Simon Benjamin, "Error-mitigated digital quantum simulation," Phys. Rev. Lett. **122**, 180501 (2019).
[48] 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 (2019).
[49] Tyler Takeshita, Nicholas C. Rubin, Zhang Jiang, Eunseok Lee, Ryan Babbush, and Jarrod R. McClean, "Increasing the representation accuracy of quantum simulations of chemistry without extra quantum resources," Phys. Rev. X **10**, 011004 (2020).
[50] Harper R. Grimsley, Sophia E. Economou, Edwin Barnes, and Nicholas J. Mayhall, "An adaptive variational algorithm for exact molecular simulations on a quantum computer," Nat. Commun. **10**, 3007 (2019).
[51] Ho Lun Tang, V. O. Shkolnikov, George S. Barron, Harper R. Grimsley, Nicholas J. Mayhall, Edwin Barnes, and Sophia E. Economou, "qubit-adapt-vqe: An adaptive algorithm for constructing hardware-efficient ansatze on a quantum processor," (2019).
[52] https://github.com/qiskit/qiskit, "Qiskit: An open-source framework for quantum computing," (2019).
[53] P. Jordan and E. Wigner, "über das paulische äquivalenzverbot," Z. Phys. **47**, 631 (1928).
[54] Jacob T. Seeley, Martin J. Richard, and Peter J. Love, "The bravyi-kitaev transformation for quantum computation of electronic structure," The Journal of Chemical Physics **137**, 224109 (2012).
[55] Sergey B. Bravyi and Alexei Yu. Kitaev, "Fermionic quantum computation," Annals of Physics **298**, 210 – 226 (2002).
[56] Kanav Setia, Sergey Bravyi, Antonio Mezzacapo, and James D. Whitfield, "Superfast encodings for fermionic

quantum simulation," Phys. Rev. Research **1**, 033033 (2019).

[57] Nikolaj Moll, Andreas Fuhrer, Peter Staar, and Ivano Tavernelli, "Optimizing qubit resources for quantum chemistry simulations in second quantization on a quantum computer," Journal of Physics A: Mathematical and Theoretical **49**, 295301 (2016).

[58] 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. Applied **11**, 044092 (2019).

[59] Michael R. Geller and Mingyu Sun, "Efficient correction of multiqubit measurement errors," arXiv preprint arXiv:2001.09980 (2020).

[60] Kathleen E. Hamilton, Tyler Kharazi, Titus Morris, Alexander J. McCaskey, Ryan S. Bennink, and Raphael C. Pooser, "Scalable quantum processor noise characterization," arXiv preprint arXiv:2006.01805 (2020).

[61] Sergey Bravyi, Sarah Sheldon, Abhinav Kandala, David C. Mckay, and Jay M. Gambetta, "Mitigating measurement errors in multi-qubit experiments," arXiv preprint arXiv:2006.14044 (2020).

[62] Kathleen E. Hamilton and Raphael C. Pooser, "Error-mitigated data-driven circuit learning on noisy quantum hardware," Quantum Machine Intelligence **2**, 115 (2020).

[63] Michael R Geller, "Rigorous measurement error correction," Quantum Science and Technology **5**, 03LT01 (2020).

[64] George S. Barron and Christopher J. Wood, "Measurement error mitigation for variational quantum algorithms," arXiv preprint arXiv:2010.08520 (2020).

[65] Kristan Temme, Sergey Bravyi, and Jay M. Gambetta, "Error mitigation for short-depth quantum circuits," Phys. Rev. Lett. **119**, 180509 (2017).

[66] Abhinav Kandala, Kristan Temme, Antonio D Córcoles, Antonio Mezzacapo, Jerry M Chow, and Jay M Gambetta, "Error mitigation extends the computational reach of a noisy quantum processor," Nature **567**, 491–495 (2019).

[67] J. T. Seeley, M. J. Richard, and P. J. Love, "The bravyi-kitaev transformation for quantum computation of electronic structure," Journal of Chemical Physics **137** (2012), Artn 224109 10.1063/1.4768229.

[68] James C. Spall, "Multivariate stochastic approximation using a simultaneous perturbation gradient approximation," IEEE transactions on automatic control **37**, 332–341 (1992).

[69] Diederik P. Kingma and Jimmy Ba, "Adam: A method for stochastic optimization," arxiv preprint arXiv: 1412.6980 (2014).

[70] M. J. D. Powell, *Advances in Optimization and Numerical Analysis*, edited by Susana Gomez and Jean-Pierre Hennart (Springer Netherlands, Dordrecht, 1994) pp. 51–67.

[71] D. R. Jones, C. D. Perttunen, and B. E. Stuckman, "Lipschitzian optimization without the lipschitz constant," Journal of Optimization Theory and Applications **79**, 157–181 (1993).

[72] M. A. Nielsen and I. L. Chuang, *Quantum Computation and Quantum Information: 10th Anniversary Edition* (Cambridge University Press, New York, 2011).