Variational Quantum Computation of Excited States

Oscar Higgott[1,2], Daochen Wang[1,3], and Stephen Brierley[1]

[1] Riverlane, 3 Charles Babbage Road, Cambridge CB3 0GT

[2] Department of Physics and Astronomy, University College London, London, WC1E 6BT

[3] Joint Center for Quantum Information and Computer Science, University of Maryland, College Park, MD 20742

July 1, 2019

The calculation of excited state energies of electronic structure Hamiltonians has many important applications, such as the calculation of optical spectra and reaction rates. While low-depth quantum algorithms, such as the variational quantum eigenvalue solver (VQE), have been used to determine ground state energies, methods for calculating excited states currently involve the implementation of high-depth controlled-unitaries or a large number of additional samples. Here we show how overlap estimation can be used to deflate eigenstates once they are found, enabling the calculation of excited state energies and their degeneracies. We propose an implementation that requires the same number of qubits as VQE and at most twice the circuit depth. Our method is robust to control errors, is compatible with error-mitigation strategies and can be implemented on near-term quantum computers.

1 Introduction

Eigenvalue problems are ubiquitous in almost all fields of science and engineering. Google's PageRank algorithm alone has had a significant impact on modern society, and at its core solves an eigenvalue problem associated with a stochastic matrix describing the World Wide Web [28]. Another important example is Principal Component Analysis (PCA) [13, 29], which has widespread applications in bioinformatics, neuroscience, image processing, and quantitative finance.

The time-independent Schrödinger equation provides yet another example of a fundamental eigenvalue problem. Its numerical solution enables properties of atoms, molecules and materials to be predicted, with far-reaching applications in materials design, drug discovery and fundamental science [38]. Characterisation of excited state energies of molecules is required to predict charge and energy transfer processes in photovoltaic

Oscar Higgott: oscar.higgott@gmail.com

materials, or to understand some chemical reactions, such as those that involve photodissociation. However, classical methods such as density functional theory are often unable to determine excited states, even for materials where ground state energy calculations are possible.

Quantum computers have the potential to solve these and other problems significantly faster than any known methods using classical computers [1, 11, 19, 36]. However many quantum algorithms will require quantum error correction, limiting their usefulness in the near future [31]. Here we study hybrid quantum-classical algorithms, which dramatically reduce the required gate depth to run and somewhat mitigate errors, by closely integrating classical and quantum subroutines [2, 9, 14, 15, 21, 23, 26, 40].

The variational quantum eigensolver (VQE), introduced in Ref. [30], is the first algorithm designed to find the lowest eigenvalue of a Hamiltonian on a near-term, non-fault-tolerant quantum computer. VQE is based on the variational principle and utilises the fact that quantum computers can store quantum states using exponentially fewer resources than required classically. VQE uses parameterised quantum circuits to prepare trial wavefunctions and compute their energy, and a classical computer to find the parameters minimising this energy. The low circuit depth of VQE has led to the hope that it may enable near-term quantum-enhanced computation.

Since its introduction, modifications have been suggested to enable VQE to find excited state energies: e.g. a folded spectrum method [30] which requires finding the expectation of the squared Hamiltonian with quadratically more terms, or symmetry-based methods which are non-systematic [23]. Such suggestions have been more recently superseded by two proposals: a method that minimises the von Neumann entropy [35] and the quantum subspace expansion method [5, 24]. However, the von Neumann entropy method ("WAVES") requires a large number of high-depth controlled-unitaries, and the quantum subspace expansion method requires a large number of additional

samples compared to VQE and introduces a new approximation.

Our algorithm extends VQE to systematically find excited states at almost no extra cost. We achieve this by adding "overlap" terms onto the optimisation function in order to exploit the fact that Hermitian matrices admit a complete set of orthogonal eigenvectors. Exploiting further the fact that VQE retains the classical parameters of ansatz states that enable their re-preparation, low-depth quantum circuits can then be readily used to calculate these overlap terms.

2 Variational quantum deflation algorithm

In VQE, the real parameters λ for the ansatz state $|\psi(\lambda)\rangle$ are classically optimised with respect to the expectation value:

$$E(\lambda) \coloneqq \langle\psi(\lambda)| H |\psi(\lambda)\rangle = \sum_j c_j \langle\psi(\lambda)| P_j |\psi(\lambda)\rangle, \quad (1)$$

of the Hamiltonian $H = \sum c_j P_j$, computed using a low-depth quantum circuit. As a result of the variational principle, finding the global minimum of $E(\lambda)$ is equivalent to finding the ground state energy of H. VQE has been implemented on many experimental platforms, and has been shown to be more resilient to control errors than the quantum phase estimation algorithm [27].

Our method extends VQE to calculate the k-th excited state by instead optimising the parameters λ_k for the ansatz state $|\psi(\lambda_k)\rangle$ such that the cost function:

$$F(\lambda_k) \coloneqq \langle\psi(\lambda_k)| H |\psi(\lambda_k)\rangle + \sum_{i=0}^{k-1} \beta_i \, |\langle\psi(\lambda_k)|\psi(\lambda_i)\rangle|^2, \quad (2)$$

is minimised. This can be seen as minimising $E(\lambda_k)$ subject to the constraint that $|\psi(\lambda_k)\rangle$ is orthogonal to the states $|\psi(\lambda_0)\rangle, ..., |\psi(\lambda_{k-1})\rangle$. In the next section, we show how choosing sufficiently large $\beta_0, ..., \beta_{k-1}$ means the minimum of $F(\lambda_k)$ is guaranteed to be the energy of the k-th state, provided that the ansatz is sufficiently expressive.

While the first term in Eq. (2) is $E(\lambda_k)$, and can be computed using the same quantum circuits as used for VQE, the second term is a sum of overlaps of the ansatz state with states 0 to $k-1$, and can be computed efficiently on a quantum computer using one of the methods given in Section 4.

Note that evaluating Eq. (2) requires knowledge of $\lambda_0, ..., \lambda_{k-1}$ and so an iterative procedure is required to calculate the k-th eigenvalue. First, λ_0 is calculated using VQE by minimising E in Eq. (1). Then, λ_1 is
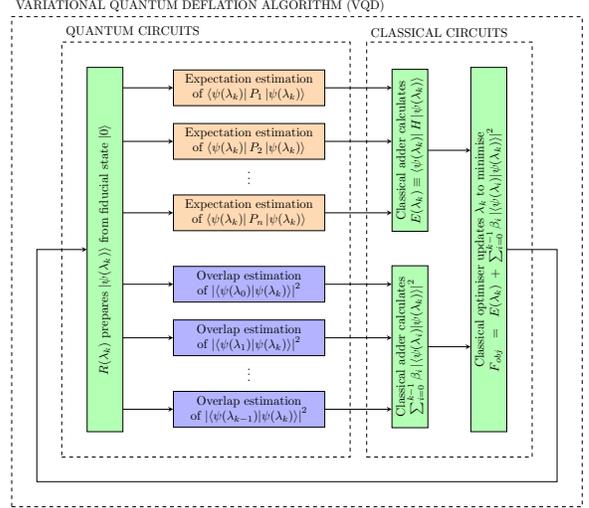

Figure 1: A schematic of our variational quantum deflation method for finding the k-th excited state of a Hamiltonian H.

calculated by minimising F in Eq. (2) for $k = 1$, after which λ_2 can be determined using the same procedure with the known λ_0 and λ_1, and so on until λ_k is determined.

A schematic of our variational quantum deflation (VQD) algorithm is shown in Fig. 1. An initial guess of λ_k is used to generate a state preparation circuit $R(\lambda_k)$ that prepares the state $|\psi(\lambda_k)\rangle$ when applied to the fiducial state $|0\rangle$. This circuit is used repeatedly to compute each of the expectation values $\langle\psi(\lambda_k)| P_j |\psi(\lambda_k)\rangle$ (see Refs. [30, 40]) and overlap terms $|\langle\psi(\lambda_k)|\psi(\lambda_i)\rangle|^2$ for $i < k$. The overlap terms are computed using circuits described in Section 4 or Appendix B or by following the method in Ref. [17].

A classical computer then uses the results of these quantum computations to calculate the objective function $F(\lambda_k)$ of Eq. (2) and update λ_k using a classical optimiser. The new λ_k is then used to prepare a new ansatz state on the quantum computer, and the whole process is repeated until some stopping criterion is reached.

As shown in Appendix A, the total number of samples $M^{(k)}$ required to measure the VQD objective function to precision ϵ when finding the kth excited state (assuming states $0 \ldots k-1$ can be perfectly prepared) is bounded above by:

$$M^{(k)} \leq \frac{1}{\epsilon^2} \left(\sum_{j=0}^{L-1} |c_j| + \frac{1}{2} \sum_{i=0}^{k-1} \beta_i \right)^2, \quad (3)$$

compared to the VQE sampling cost of $M \leq \frac{1}{\epsilon^2} \left(\sum_{j=0}^{L-1} |c_j| \right)^2$. For well-chosen β_i, we expect this additional sampling cost relative to VQE to be very small,

as explained in more detail in Appendix A.

3 Overlap weighting

An equivalent viewpoint of our optimisation procedure is that we are finding the ground state of the effective Hamiltonian at stage k:

$$H_k := H + \sum_{i=0}^{k-1} \beta_i \left| i \right\rangle \left\langle i \right|, \qquad (4)$$

where $\left| i \right\rangle$ is the (previously found) i-th eigenstate of H with energy $E_i := \left\langle i \right| H \left| i \right\rangle$ [1]. It can be easily verified that for an arbitrary state $\left| \psi \right\rangle := \sum a_i \left| i \right\rangle$:

$$\left\langle \psi \right| H_k \left| \psi \right\rangle = \sum_{i=0}^{k-1} |a_i|^2 (E_i + \beta_i) + \sum_{i=k}^{d-1} |a_i|^2 E_i,$$

where d is the total number of eigenvectors of H.

Therefore, if the ansatz is sufficiently powerful, then to guarantee a minimum at E_k, it suffices to choose $\beta_i > E_k - E_i$. Since $\Delta := E_{d-1} - E_0 \geq E_k - E_i$, it suffices to possess an accurate estimate of Δ, e.g. by using VQE to find E_0 and then E_{d-1} (using the Hamiltonian $-H$ to find the latter). When we readily have a specification of $H = \sum c_j P_j$ as a linear combination of Pauli matrices, e.g. when H is the electronic structure Hamiltonian, then we have the upper bound $\Delta \leq 2\|H\| \leq 2\sum |c_j|$. In this case, we can readily choose β_i to guarantee the validity of our procedure.

Choosing valid β_i can also be self-correcting. For example, if we incorrectly chose $\beta_i = \gamma - E_i \leq E_k - E_i$ for all i, we will discover that we have set β_i too small since we will eventually find a minimum at $F(\lambda_k) = \gamma$. However, by repeating the algorithm with a larger γ until an energy strictly less than γ is found (doubling γ each time, say), we can pick a large enough γ after $O(\log(E_k - E_0))$ runs of the algorithm.

4 Low-depth implementations

A low-depth method for overlap estimation, proposed in Ref. [12], can be seen by writing the overlap $|\langle \psi(\lambda_i)|\psi(\lambda_k)\rangle|^2$ as $|\langle 0| R(\lambda_i)^\dagger R(\lambda_k) |0\rangle|^2$. We can prepare the state $R(\lambda_i)^\dagger R(\lambda_k) |0\rangle$ using the trial state preparation circuit followed by the inverse of the preparation circuit for the i-th previously-computed state. The overlap is then estimated to precision ϵ by the fraction of all-zero bitstrings when measuring this state $O(1/\epsilon^2)$ times in the computational basis.

[1] We assume these are true eigenstates with possibly non-distinct energies.

This method requires knowing the inverse of the preparation circuit for each previously-computed state, $R(\lambda_i)^\dagger$. While this inverse is often known in theory by inverting gates in a decomposition of the original preparation circuit, device errors may mean that the implementation is inaccurate in practice. If we define λ_i^* to be the optimal parameters originally found to prepare the i-th state $R(\lambda_i^*) |0\rangle$ using VQD, then its inverse can be found by fixing λ_i^* and varying the trial state parameters λ_i such that the overlap $|\langle 0| R(\lambda_i)^\dagger R(\lambda_i^*) |0\rangle|^2$ is maximised. This technique enables VQD to retain the robustness to control errors that is characteristic of VQE [27].

This implementation of VQD requires the same number of qubits as VQE and around twice the circuit depth. In Appendix B, we describe an alternative method which uses the destructive SWAP test and requires almost the same circuit depth as VQE but twice the number of qubits. If a larger gate-depth is available, then α-QPE [40] can be used to reduce the total runtime of overlap estimation from $O(\frac{1}{\epsilon^2})$ up to $O(\frac{1}{\epsilon})$.
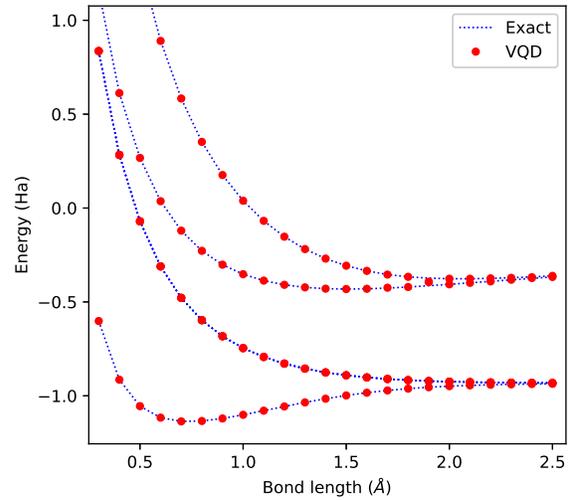

Figure 2: All ground and excited state energy levels of H_2 in the STO-3G basis, calculated using exact diagonalisation (blue dotted line) and our variational quantum deflation (VQD) method (red filled circles) over a range of internuclear separations.

5 Numerical simulation: H_2

We simulated VQD on H_2 in the STO-3G basis for a range of internuclear separations and compared it to exact diagonalisation, as shown in Fig. 2. Using $\beta_i = 3$ Ha for all i and a generalised unitary coupled cluster singles and doubles (UCCGSD) ansatz, the median error of our method relative to exact diagonalisation is less than

4×10^{-6} Ha for all energy levels, significantly better than chemical accuracy of 1.6×10^{-3} Ha (the precision required to determine reaction rates to within an order of magnitude at room temperature using the Eyring equation [8]). Our method finds all 6 eigenstates systematically, including all those in the 3-dimensional degenerate subspace spanned by the 1st, 2nd and 3rd excited states. The ability to find degenerate states is another key advantage of our method; the folded spectrum and WAVES methods rely on the energies of states to differentiate between them and have no systematic way of determining the degeneracy of the eigenvalues. Further discussion of our simulation, including optimiser and ansatz used, can be found in Appendix C.

6 Error accumulation

In general, we cannot assume perfect state preparation for states $i < k$. Suppose a state $|\tilde{\psi_0}\rangle$ (with energy $\tilde{E}_0$) is prepared instead of the true ground state $|\psi_0\rangle$ such that $\left|\langle\tilde{\psi_0}|\psi_0\rangle\right|^2 = 1 - \epsilon_0$, leading to an error in the ground state energy $\epsilon'_0 = \tilde{E}_0 - E_0 = O(2\epsilon_0||H||)$. If we now use this ground state estimate along with VQD to find the first excited state $|\psi_1\rangle$ using a new trial state $|\tilde{\psi_1}\rangle$, the lowest-energy state of the deflated Hamiltonian no longer corresponds to the exact excited state energy E_1.

The inexact deflated Hamiltonian is now given by $\tilde{H}_1 = H + \beta_0 |\tilde{\psi_0}\rangle\langle\tilde{\psi_0}|$ and, to assess the accumulation of errors, we wish to find upper and lower bounds for $\min_{\tilde{\psi_1}}[\langle\tilde{\psi_1}| \tilde{H}_1 |\tilde{\psi_1}\rangle]$.

In Appendix D we show that, provided we set $\beta_0 > \frac{E_1 - E_0}{1 - \epsilon_0}$, the ground state energy of the inexact deflated Hamiltonian is bounded by terms linear in ϵ_0:

$$E_1 - O((E_1 - E_0)\epsilon_0) \leq \min_\psi \langle\psi| \tilde{H}_1 |\psi\rangle \leq E_1 + \beta_0\epsilon_0. \quad (5)$$

In reality we will not find the exact ground state of the deflated Hamiltonian $\tilde{H}_1$ and will incur an additional error ϵ_1 as was the case for the ground state of the original Hamiltonian H. However, provided $\epsilon_1 \approx \epsilon_0$, our total error $\epsilon'_1 = O(\epsilon_0\beta_0 + 2\epsilon_1||H||)$ in the energy is still linear in our original ground state error. An alternative analysis of error accumulation is provided by Lee et al. [18].

We analysed this accumulation of errors further through numerical simulations of VQD in the presence of sampling error, shown in Fig. 3. We analysed three different sampling rates: $M = 10^6$, 10^7 and 10^8 samples per Hamiltonian subterm and overlap term, running 225 simulations of VQD (with random initial parameters) for each of these three scenarios. Of these runs, $\sim 20\%$ of the simulations found the eigenstates in the incorrect order and were discarded for consistency
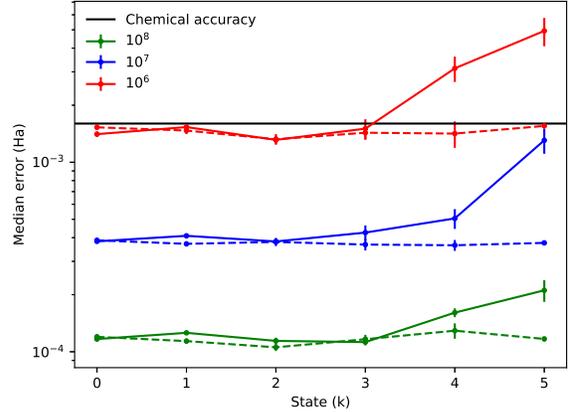

Figure 3: Median error using VQD to determine each energy level k in the spectrum of H_2 at bond distance (0.7414 Å) in the STO-3G basis. Red, blue and green lines show results using 10^6, 10^7 and 10^8 samples (per Hamiltonian subterm and overlap term) respectively. For the solid lines, the standard VQD algorithm was used, whereas for dashed lines states $i < k$ in H_k were computed exactly for comparison. Chemical accuracy (1.6×10^{-3} Ha) is also shown for reference (black solid line). Error bars show 1σ standard errors for the median estimates (calculated using bootstrap resampling [6]).

in the analysis. The median errors for the remaining ~ 180 runs for each state k are shown in Fig. 3. For comparison, we also simulated 130 runs (dashed lines) using 'exact' states $i < k$ in H_k ($< 10^{-7}$ energy error in each state $i < k$). For all three sampling rates, the median error in the first excited state is similar in magnitude to the error in the ground state, as expected from our analysis earlier in this section and in Appendix D. Furthermore, the median errors for all states $k < 4$ are very similar (for a given M) to the errors when using an exact H_k, and are all below chemical accuracy, demonstrating that error accumulation is negligible for these states. For $k = 4$ and $k = 5$ the accumulated error is substantially higher than the error using an exact H_k, however, showing that VQD is most effective for low-lying states. Achieving chemical accuracy for $k = 5$ requires 10^7 samples, instead of 10^6 for $0 < k < 4$.

One way to address this accumulation of errors within VQD to find higher excited states may be to use the alternative effective Hamiltonians discussed in Section 7. Another solution is to use a hybrid approach, using VQD instead of excitation operators in the WAVES protocol [35]. Here, VQD may provide a more effective method of approximating excited states than the excitation operators proposed in WAVES, whereas the von-Neumann entropy "eigenstate witness" used in WAVES does not have the same problem of error accumulation,

and could help refine the energy estimate. Both of these alternative approaches require a larger gate depth than the version of VQD we have studied here, but may be a good approach to finding higher excited states in the era of fault-tolerant quantum computing.

7 Choice of effective Hamiltonian

The form of our effective Hamiltonian in Eq. (4) is only one choice within the broad category of deflation methods. Such methods are typically employed to find eigenvalues and eigenvectors of positive semi-definite matrices, often covariance matrices in the context of PCA, starting from the largest eigenvalues.

To make direct use of deflation methods for positive semi-definite matrices, note that the Hamiltonian $H' := -H + E'$ for some $E' \geq E_{d-1}$, e.g. $E' = \|H\|$, is positive semi-definite. Under this transformation, we find that Hotelling's deflation corresponds to our method and would set $\beta_i = E' - E_i$ in Eq. (4).

Other deflation methods exist such as projection deflation or Schur complement deflation which are designed to address the problem of not obtaining true eigenstates at each stage. These two methods, in contrast to Hotelling's, ensure that the true ground state of the effective Hamiltonian at each stage does not overlap with the previously found eigenstate estimate irrespective of its accuracy. Empirically, these two methods have been found to perform better than Hotelling's in the context of PCA on some datasets [20].

For example, in projection deflation, the effective Hamiltonian at stage k is defined as:

$$H_k = A_k^\dagger (H - E') A_k, \tag{6}$$

where:

$$A_k := \prod_{i=0}^{k-1} (1 - |i\rangle \langle i|) \approx 1 - \sum_{i=0}^{k-1} |i\rangle \langle i|, \tag{7}$$

and the last approximation holds when the previously found eigenvectors $|i\rangle$ are truly orthogonal.

With this approximation, writing H again as a linear combination of Pauli matrices P_j, the value of $\langle\psi| H_k |\psi\rangle$ for an ansatz $|\psi\rangle$ is a linear combination of terms of forms: $\langle\psi| P_j |\psi\rangle$, $|\langle\psi | i\rangle|^2$ as in Hotelling's deflation, but now additionally $\langle\psi | i\rangle$, $\langle\psi| P_j |i\rangle$ and $\langle i| P_j |l\rangle$ (for $i, l < k$). Without this approximation, we also need to calculate $\langle i | l\rangle$. The important point now is that all these additional terms can still be quantum computed, e.g. following the method in Ref. [17].

8 Discussion

We have introduced a new method–variational quantum deflation (VQD)–for calculating low-lying excited state energies of quantum systems using a quantum computer. Our method requires the same number of qubits as the variational quantum eigensolver (VQE) for ground state methods, at most twice the maximum circuit depth (for any given ansatz) and a negligible increase in the number of required measurements. By contrast, existing methods for quantum computing excited states require a large overhead in resources compared to ground state methods.

While we used a Nelder-Mead optimiser and UCCSD ansatz in our simulation of molecular Hydrogen here, we note that many other optimisers and ansatz circuits can also be used for VQD. After the first version of this paper was released, interesting work by Jones et al. compared the use of two different optimisation methods as applied to our protocol to calculate the spectrum of a Lithium Hydride molecule [15]. More recently, work by Lee et al. showed that using a multi-determinental reference state or their k-UpCCGSD ansatz can improve the precision of finding the first excited state of N_2 using VQD [18]. Further work could include numerical analysis of different optimisers and ansatz circuits for use within VQD in the presence of noise, as well as the effectiveness of the alternative effective Hamiltonians presented in Section 7.

Given its low resource requirements and compatibility with error-mitigation techniques, we hope that VQD may enable the quantum-enhanced computation of excited state energies in the near-future.

A Sampling cost

In VQE, the variance ϵ^2 in the energy expectation value $\langle H \rangle$ after using M_j samples for the measurement of each subterm $\langle P_j \rangle$ in the Hamiltonian $H = \sum c_j P_j$ is bounded by [30, 33]:

$$\epsilon^2 = \sum_{j=0}^{L-1} \frac{c_j^2 \sigma_j^2}{M_j} \tag{8}$$

$$= \sum_{j=0}^{L-1} \frac{c_j^2 (1 - \langle P_j \rangle^2)}{M_j} \leq \sum_{j=0}^{L-1} \frac{c_j^2}{M_j}. \tag{9}$$

where $\sigma_j^2 = \mathrm{Var}\left[\langle P_j \rangle\right] = \langle P_j^2 \rangle - \langle P_j \rangle^2$ is the intrinsic variance of the projective measurement of $\langle P_j \rangle$. Using the method of Lagrange multipliers, Rubin et al. [33] showed that the optimal choice of M_j to minimise the total number of samples $M = \sum_j M_j$ used to achieve

precision ϵ is:

$$M_j = \frac{1}{\epsilon^2} |c_j| \sigma_j \sum_{i=0}^{L-1} |c_i| \sigma_i, \tag{10}$$

which leads to a total number of samples

$$M = \frac{1}{\epsilon^2} \left(\sum_{j=0}^{L-1} |c_j| \sigma_j \right)^2 \leq \frac{1}{\epsilon^2} \left(\sum_{j=0}^{L-1} |c_j| \right)^2. \tag{11}$$

Assuming perfect state preparation for states $i < k$ in VQD, we find that the variance of the energy expectation value $\langle H_k \rangle$ of the deflated Hamiltonian H_k is instead given by:

$$\epsilon^2 = \sum_{j=0}^{L-1} \frac{c_j^2 \sigma_j^2}{M_j^k} + \sum_{i=0}^{k-1} \frac{\beta_i^2 \tilde{\sigma}_i^2}{\tilde{M}_i^k} \tag{12}$$

$$\leq \sum_{j=0}^{L-1} \frac{c_j^2}{M_j^k} + \sum_{i=0}^{k-1} \frac{\beta_i^2}{4\tilde{M}_i^k} \tag{13}$$

where M_j^k is the number of samples used for measuring $\langle P_j \rangle$, $\tilde{M}_i^k$ is the number of samples used to estimate the overlap of the ansatz with the ith previously found state and $\tilde{\sigma}_i^2 = |\langle i|k \rangle|^2 \left(1 - |\langle i|k \rangle|^2\right)$ is the intrinsic variance of this overlap measurement. From a straightforward extension of the Lagrange multiplier approach used by Rubin et al. [33], we now find the optimal M_j^k and $\tilde{M}_i^k$ for the deflated Hamiltonian H_k to be:

$$M_j^k = \frac{1}{\epsilon^2} |c_j| \sigma_j \left(\sum_{l=0}^{L-1} |c_l| \sigma_l + \sum_{i=0}^{k-1} \beta_i \tilde{\sigma}_i \right), \tag{14}$$

$$\tilde{M}_i^k = \frac{1}{\epsilon^2} \beta_i \tilde{\sigma}_i \left(\sum_{j=0}^{L-1} |c_j| \sigma_j + \sum_{l=0}^{k-1} \beta_l \tilde{\sigma}_l \right). \tag{15}$$

This leads to a total number of samples $M^{(k)}$ given by:

$$M^{(k)} = \sum_{j=0}^{L-1} M_j^k + \sum_{i=0}^{k-1} \tilde{M}_i^k \tag{16}$$

$$= \frac{1}{\epsilon^2} \left(\sum_{j=0}^{L-1} |c_j| \sigma_j + \sum_{i=0}^{k-1} \beta_i \tilde{\sigma}_i \right)^2 \tag{17}$$

$$\leq \frac{1}{\epsilon^2} \left(\sum_{j=0}^{L-1} |c_j| + \frac{1}{2} \sum_{i=0}^{k-1} \beta_i \right)^2. \tag{18}$$

From a comparison of $M^{(k)}$ with M we expect the sampling overhead of VQD relative to VQE to be very small, since the sum of the $L = O(N^4)$ Hamiltonian coefficients will likely be far larger than the sum of well-chosen weights β_i for low-lying excited states in practice. Furthermore, the variances $\tilde{\sigma}_i^2$ tend to zero at convergence. However, if we require precision ϵ throughout

the optimisation rather than just at convergence, and if we choose $\beta_i = 2 \sum_{j=0}^{L-1} |c_j|$ since this always guarantees β_i is large enough, then we find $M^{(k)} = (1+k)^2 M$ (where we have used the upper bounds for both $M^{(k)}$ and M).

B Destructive SWAP test

The SWAP test enables the overlap $|\langle \phi | \psi \rangle|^2$ of two states $|\psi\rangle$ and $|\phi\rangle$ to be determined to precision ϵ using $O(1/\epsilon^2)$ repeated measurements after applying a circuit to a quantum register in the state $|\psi\rangle \otimes |\phi\rangle$. While the original SWAP test acting on two N-qubit states required an ancilla and a controlled-SWAP gate, leading to a $2N + 1$-qubit circuit with depth $O(N)$, it was shown in Refs. [4, 10] that the same outcome distribution can be attained more efficiently without an ancilla, using parallel Bell-basis measurements and classical logic. This so-called "destructive SWAP test" (shown in Fig. 4) requires just $2N$ qubits and depth $O(1)$, achieving significant savings compared to the original SWAP test.
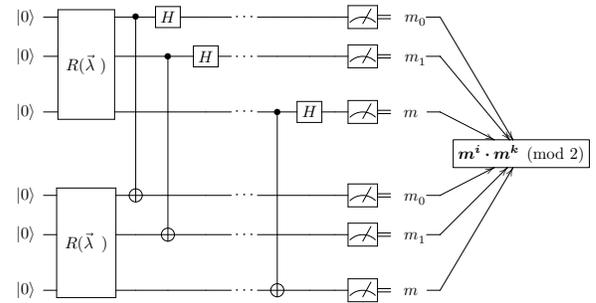

Figure 4: The N-qubit generalisation of the destructive SWAP test as applied to two ansatz states $|\psi(\lambda_i)\rangle$ and $|\psi(\lambda_k)\rangle$, prepared using state preparation circuits $R(\lambda_i)$ and $R(\lambda_k)$ respectively.

If the ansatz used can be implemented on a linear chain of qubits with nearest neighbour connectivity, e.g. parameterised adiabatic state preparation using the fermionic SWAP network Trotter step [16], then the SWAP test to compare two ansatz states can be implemented on a $N \times 2$ nearest-neighbour grid quantum computer architecture with a depth-one circuit that is subgraph isomorphic to the architecture (i.e. no routing of quantum information required). This implementation makes the assumption that the same ansatz state can be prepared with the same parameters on two separate registers of qubits. If this cannot be assumed (e.g. if qubit errors are inhomogeneous), then the SWAP test can be used to "copy" the state from the first register to the second register, by maximising the overlap of the

two states, with the parameters of the state on the second register allowed to vary. This technique allows the SWAP test implementation of VQD to maintain robustness to control errors.

C   Methods for numerical simulation

The standard UCCSD ansatz [32] is defined relative to a reference state $|\psi_0\rangle$ by:

$$|\psi\rangle = e^{T-T^\dagger} |\psi_0\rangle ,$$

where $T := T_1 + T_2$ with:

$$T_1 := \sum_{\substack{i\in\text{occ}\\l\in\text{vir}}} t_i^l a_m^\dagger a_i,$$

$$T_2 := \sum_{\substack{i,j\in\text{occ}\\l,k\in\text{vir}}} t_{ij}^{lk} a_l^\dagger a_k^\dagger a_i a_j,$$

for some parameters t_i^l, $t_{ij}^{lk} \in \mathbb{R}$ and occ and vir are the sets of occupied and virtual orbitals of $|\psi_0\rangle$.

We instead use a generalised unitary coupled cluster ansatz (UCCGSD) with $|\psi_0\rangle = |\text{HF}\rangle$ set to the Hartree-Fock state but with the cluster operator $T = T_1 + T_2$ now using the definitions:

$$T_1 := \sum_{pq} t_p^q a_q^\dagger a_p,$$

$$T_2 := \sum_{pqrs} t_{pq}^{rs} a_r^\dagger a_s^\dagger a_p a_q,$$

where p, q, r, s can now index any orbital (irrespective of its occupation in the reference state). A variant of this ansatz was suggested by McClean et al. [23] in the context of VQE, and UCCGSD has since been investigated numerically [18, 41]. Lee et al. found UCCGSD to perform significantly better than UCCSD in VQE for a number of small molecules [18].

Since we are only interested in the parameterisation of $T - T^\dagger$, and fermionic operators obey the anti-commutation relations:

$$\{a_j, a_k\} = 0, \ \{a_j^\dagger, a_k^\dagger\} = 0, \ \{a_j, a_k^\dagger\} = \delta_j^k,$$

it can be directly verified that there are only 6 and 3 independent parameters for T_1 (e.g. t_0^1, t_0^2, t_0^3, t_1^2, t_1^3, t_2^3) and T_2 (e.g. t_{01}^{23}, t_{02}^{13}, t_{03}^{12}) respectively.

The results in Fig. 2 were simulated using ProjectQ and FermiLib [25, 37]. A tolerance of 10^{-2} was used with a Nelder-Mead optimiser (xatol=fatol=10^{-2}, as implemented in the scipy Python scientific library), and the best of two consecutive (randomly initialised) runs was used for each bond length and energy level. We note

that other optimisers, such as LGO, have been shown to offer improved performance in VQE [23], and possible further work includes analysis of alternative optimisation strategies in the context VQD.

While we initialised the UCCGSD parameters randomly in this work, choosing a good initial guess for the parameters can significantly reduce the number of iterations required for the optimiser to converge. For ground state VQE problems, second order Møller-Plesset perturbation theory (MP2) has previously been proposed as a UCC ansatz initialisation method [32]. Another method, for either ground or excited states, initialises a UCC ansatz with optimised expectation values that are classically estimated using a truncated BCH expansion [18].

We also note that the overlap terms $|\langle\psi(\lambda_k)|\psi(\lambda_i)\rangle|^2$ in Eq. (2) of the state k with a known state i are similar to the overlap terms $|\langle\psi(\lambda_s^t)|\psi(\lambda_i)\rangle|^2$ of the same known state i with another previously-computed state s (where $i < s < k$) in the t-th iteration of the VQD optimisation procedure used to compute that state. It may therefore be advantageous to cache the outputs of these $|\langle\psi(\lambda_s^t)|\psi(\lambda_i)\rangle|^2$ terms, and use them to inform and improve the optimisation procedure for the k-th state, hopefully reducing the number of optimisation steps and quantum circuits required.

D   Bounds for error accumulation

In Section 6 we stated that, to assess the accumulation of errors, we would like to find upper and lower bounds for the ground state energy $\min_\psi[\langle\psi|\tilde{H}_1|\psi\rangle]$ of the inexact deflated Hamiltonian $\tilde{H}_1 = H + \beta_0 |\tilde{\psi}_0\rangle\langle\tilde{\psi}_0|$, where $\tilde{\psi}_0$ is the (inexact) estimate of the ground state found in the first iteration of VQD.

Using the same notation as in Section 6, and writing states in the eigenbasis of the Hamiltonian, $|\tilde{\psi}_0\rangle = \sum_{i=0}^{d-1} a_i |\psi_i\rangle$ and $|\tilde{\psi}_1\rangle = \sum_{i=0}^{d-1} b_i |\psi_i\rangle$, an $O(\beta_0\epsilon_0)$ upper bound is given straightforwardly by $\langle\psi_1|\tilde{H}_1|\psi_1\rangle \le E_1 + \beta_0\epsilon_0$. Writing $\vec{a_1} := (a_1,\ldots,a_{d-1})$ and $\vec{b_1} := (b_1,\ldots,b_{d-1})$ for compactness, we find the lower bound to be:

$$\langle\tilde{\psi}_1|\tilde{H}_1|\tilde{\psi}_1\rangle \tag{19}$$

$$= |b_0|^2 E_0 + \sum_{i=1}^{d-1} |b_i|^2 E_i + \beta_0 |a_0^* b_0 + \langle\vec{a_1}, \vec{b_1}\rangle|^2 \tag{20}$$

$$= |b_0|^2(E_0 + \beta_0|a_0|^2) + \sum_{i=1}^{d-1} |b_i|^2 E_i$$
$$+ \beta_0(2\Re(a_0^* b_0 \langle\vec{a_1}, \vec{b_1}\rangle^*) + |\langle\vec{a_1}, \vec{b_1}\rangle|^2) \tag{21}$$
$$\ge |b_0|^2(E_0 + \beta_0(1 - \epsilon_0)) - |b_0|(2\beta_0\sqrt{\epsilon_0})$$

$$+ \min_{\vec{b_1}} \sum_{i=1}^{d-1} |b_i|^2 E_i \qquad (22)$$

$$\geq |b_0|^2 (\beta_0(1 - \epsilon_0) - (E_1 - E_0))$$
$$- |b_0|(2\beta_0\sqrt{\epsilon_0}) + E_1 \qquad (23)$$

$$\geq E_1 - \epsilon_0 \frac{\beta_0^2}{\beta_0(1 - \epsilon_0) - (E_1 - E_0)}. \qquad (24)$$

where the first inequality is Cauchy-Schwarz, the second inequality follows from $\sum_{i=0}^{d-1} |b_i|^2 = 1$, and the third inequality follows by minimising a quadratic over $|b_0|$ (assuming $\beta_0 > \frac{E_1 - E_0}{1 - \epsilon_0}$). From the Taylor series expansion in ϵ_0 of the second inequality we find:

$$\langle \tilde{\psi}_1| \tilde{H}_1 |\tilde{\psi}_1\rangle \geq E_1 - \frac{\beta_0(E_1 - E_0)}{\beta_0 - (E_1 - E_0)}\epsilon_0 + O(\epsilon_0^2), \quad (25)$$

from which it is clear that, for any fixed $\beta_0 > \frac{E_1 - E_0}{1 - \epsilon_0}$, we have a lower bound of:

$$\min_{\tilde{\psi}_1} \langle \tilde{\psi}_1| \tilde{H}_1 |\tilde{\psi}_1\rangle \geq E_1 - O((E_1 - E_0)\epsilon_0). \qquad (26)$$

E Symmetry constraints

It is often the case that the Hilbert space of the Hamiltonian being considered is larger than the Hilbert space relevant to the particular problem of interest. For example, consider the electronic structure Hamiltonian in second quantised form:

$$H = \sum_{ij} h_{ij} a_i^\dagger a_j + \sum_{ijkl} h_{ijkl} a_i^\dagger a_j^\dagger a_k a_l, \qquad (27)$$

where $a_i^\dagger$ and a_i are the fermionic creation and annihilation operators for an electron in the i-th spin orbital, and where the coefficients h_{ij} and h_{ijkl} denote the one- and two-electron integrals, respectively. After the Hamiltonian is transformed through the Jordan-Wigner or Bravyi-Kitaev transformation, converting creation and annihilation operators into qubit operators, the dimension of the Hilbert space remains 2^N, where N is the number of spin orbitals. However, if one is interested only in states with a particular symmetry, the dimension of the Hilbert space restricted only to these states can be much smaller, e.g. $\binom{N}{\eta} = O(N^\eta)$ instead of 2^N if only η-electron states are of interest.

If we wish to apply VQD to find excited states of a molecular Hamiltonian with a particular symmetry, it is necessary that the ansatz state for a desired excited state, at the global minimum of Eq. (2), be contained entirely within the restricted Hilbert space of interest. One way of ensuring this is to use an ansatz that always conserves the correct symmetry. For example, the fermionic unitary coupled cluster ansatz we use in Section 5 conserves the desired number of electrons ($\eta = 2$) of neutral molecular Hydrogen for all input parameters.

Alternatively, penalty terms can be included in the objective function such that the ansatz state has the desired symmetry at the global minimum of the objective function [23, 34]. This leads to a modified objective function:

$$F_C(\lambda_k) := F(\lambda_k) + \sum_i \mu_i \big[\langle \psi(\lambda_k)| \hat{C}_i |\psi(\lambda_k)\rangle - c_i \big]^2, \qquad (28)$$

where $\hat{C}_i$ are symmetry constraining operators (e.g. $\hat{N}_e$, $\hat{S}^2$, $\hat{S}_z$) and c_i are constants corresponding to their desired expectation values.

Clearly, by incorporating any of these techniques, we can find the excited states of a Hamiltonian constrained to any particular symmetry of interest.

F Error mitigation

In Refs. [3, 22, 34], an error-mitigating post-processing procedure was introduced that uses the operators $\hat{C}_i$ (defined in Appendix E) to detect and discard all measurements that violate a required symmetry for energy expectation circuits in VQE-type algorithms. This procedure can produce more accurate expectation values in the presence of bit-flip errors and some combinations of two-qubit errors.

After the first version of this paper was released, Ref. [15] incorporated our VQD technique to calculate excited states using imaginary time evolution. The authors also proposed a method to detect symmetry-breaking errors when using the ancilla-based SWAP-test, by performing symmetry measurements on the ansatz registers while measuring the overlap with the ancilla. However, using the low-depth overlap estimation circuit given in Section 4, we can detect and discard any error that does not commute with a symmetry operator $\hat{C}_i$ using classical post-processing alone, provided that $\hat{C}_i$ is diagonal in the computational basis and commutes with the ansatz. In the Jordan-Wigner and Bravyi-Kitaev encodings, the operators for the number of electrons $\hat{N}_e$, spin up electrons $\hat{N}_\uparrow$ and spin down electrons $\hat{N}_\downarrow$ are diagonal in the computational basis, allowing these quantities to be computed classically in post-processing for both encodings. For example, starting from $|0\rangle$, the UCC ansatz is prepared by $R_{\text{UCC}}(\lambda) = V(\lambda)R_{\text{HF}}$, where R_{HF} prepares the Hartree-Fock state $|\text{HF}\rangle$ and $V := e^{T - T^\dagger}$ is the UCC operator. Now, rather than measuring the fraction of all-zero bitstrings after performing $R_{\text{UCC}}(\lambda_i)^\dagger R_{\text{UCC}}(\lambda_k) |0\rangle = R_{\text{HF}}^\dagger V(\lambda_i)^\dagger V(\lambda_k) R_{\text{HF}} |0\rangle$, we can instead measure the

fraction of bitstrings corresponding to $|\text{HF}\rangle$ after performing $V(\lambda_i)^\dagger V(\lambda_k) R_{HF} |0\rangle$. Since V conserves electron number, we know that all measured bitstrings that do not correspond to the correct electron number can be discarded as per the post-processing procedure. Therefore, this method for error-mitigated overlap estimation is more efficient than the ancilla-based method proposed in Ref. [15] if $\hat{C}_i$ is diagonal in the computational basis. We also note that the error-mitigation techniques proposed in Refs. [7, 39] can be readily applied to our algorithm.

References

[1] Alán Aspuru-Guzik, Anthony D. Dutoi, Peter J. Love, and Martin Head-Gordon. Simulated quantum computation of molecular energies. *Science*, 309(5741):1704–1707, 2005. ISSN 0036-8075. DOI: 10.1126/science.1113479. URL https://doi.org/10.1126/science.1113479.

[2] 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):45, 2019. DOI: 10.1038/s41534-019-0157-8. URL https://doi.org/10.1038/s41534-019-0157-8.

[3] X. Bonet-Monroig, R. Sagastizabal, M. Singh, and T. E. O'Brien. Low-cost error mitigation by symmetry verification. *Phys. Rev. A*, 98:062339, Dec 2018. DOI: 10.1103/PhysRevA.98.062339. URL https://doi.org/10.1103/PhysRevA.98.062339.

[4] Lukasz Cincio, Yiğit Subaşı, Andrew T Sornborger, and Patrick J Coles. Learning the quantum algorithm for state overlap. *New Journal of Physics*, 20(11):113022, nov 2018. DOI: 10.1088/1367-2630/aae94a. URL https://doi.org/10.1088%2F1367-2630%2Faae94a.

[5] 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, Feb 2018. DOI: 10.1103/PhysRevX.8.011021. URL https://doi.org/10.1103/PhysRevX.8.011021.

[6] B. Efron. Bootstrap methods: Another look at the jackknife. *Ann. Statist.*, 7(1):1–26, 01 1979. DOI: 10.1214/aos/1176344552. URL https://doi.org/10.1214/aos/1176344552.

[7] Suguru Endo, Simon C. Benjamin, and Ying Li. Practical quantum error mitigation for near-future

applications. *Phys. Rev. X*, 8:031027, Jul 2018. DOI: 10.1103/PhysRevX.8.031027. URL https://doi.org/10.1103/PhysRevX.8.031027.

[8] Henry Eyring. The activated complex in chemical reactions. *The Journal of Chemical Physics*, 3 (2):107–115, 1935. DOI: 10.1063/1.1749604. URL https://doi.org/10.1063/1.1749604.

[9] Edward Farhi, Jeffrey Goldstone, and Sam Gutmann. A quantum approximate optimization algorithm. *arXiv preprint arXiv:1411.4028*, 2014. URL https://arxiv.org/abs/1411.4028.

[10] Juan Carlos Garcia-Escartin and Pedro Chamorro-Posada. The swap test and the Hong-Ou-Mandel effect are equivalent. *Phys. Rev. A*, 87:052330, May 2013. DOI: 10.1103/PhysRevA.87.052330. URL https://doi.org/10.1103/PhysRevA.87.052330.

[11] Lov K. Grover. A fast quantum mechanical algorithm for database search. In *Proceedings of the Twenty-eighth Annual ACM Symposium on Theory of Computing*, STOC '96, pages 212–219, New York, NY, USA, 1996. ACM. ISBN 0-89791-785-5. DOI: 10.1145/237814.237866. URL https://doi.org/10.1145/237814.237866.

[12] Vojtěch Havlíček, Antonio D Córcoles, Kristan Temme, Aram W Harrow, Abhinav Kandala, Jerry M Chow, and Jay M Gambetta. Supervised learning with quantum-enhanced feature spaces. *Nature*, 567(7747):209, 2019. DOI: 10.1038/s41586-019-0980-2. URL https://doi.org/10.1038/s41586-019-0980-2.

[13] Harold Hotelling. Analysis of a complex of statistical variables into principal components. *Journal of educational psychology*, 24(6):417, 1933. DOI: 10.1037/h0071325. URL https://doi.org/10.1037/h0071325.

[14] Peter D Johnson, Jonathan Romero, Jonathan Olson, Yudong Cao, and Alán Aspuru-Guzik. QVECTOR: an algorithm for device-tailored quantum error correction. *arXiv preprint arXiv:1711.02249*, 2017. URL https://arxiv.org/abs/1711.02249.

[15] Tyson Jones, Suguru Endo, Sam McArdle, Xiao Yuan, and Simon C. Benjamin. Variational quantum algorithms for discovering hamiltonian spectra. *Phys. Rev. A*, 99:062304, Jun 2019. DOI: 10.1103/PhysRevA.99.062304. URL https://doi.org/10.1103/PhysRevA.99.062304.

[16] 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, Mar 2018. DOI: 10.1103/Phys-

RevLett.120.110501. URL https://doi.org/10.1103/PhysRevLett.120.110501.

[17] Emanuel Knill, Gerardo Ortiz, and Rolando D. Somma. Optimal quantum measurements of expectation values of observables. *Phys. Rev. A*, 75:012328, Jan 2007. DOI: 10.1103/PhysRevA.75.012328. URL https://doi.org/10.1103/PhysRevA.75.012328.

[18] Joonho Lee, William J Huggins, Martin Head-Gordon, and K Birgitta Whaley. Generalized unitary coupled cluster wavefunctions for quantum computation. *Journal of chemical theory and computation*, 2018. DOI: 10.1021/acs.jctc.8b01004. URL https://doi.org/10.1021/acs.jctc.8b01004.

[19] Seth Lloyd, Masoud Mohseni, and Patrick Rebentrost. Quantum principal component analysis. *Nature Physics*, 10(9):631, 2014. DOI: 10.1038/nphys3029. URL https://doi.org/10.1038/nphys3029.

[20] Lester W. Mackey. Deflation methods for sparse PCA. pages 1017–1024, 2009. URL http://papers.nips.cc/paper/3575-deflation-methods-for-sparse-pca.pdf.

[21] Sam McArdle, Suguru Endo, Ying Li, Simon Benjamin, and Xiao Yuan. Variational quantum simulation of imaginary time evolution with applications in chemistry and beyond. *arXiv preprint arXiv:1804.03023*, 2018. URL https://arxiv.org/abs/1804.03023.

[22] Sam McArdle, Xiao Yuan, and Simon Benjamin. Error-mitigated digital quantum simulation. *Phys. Rev. Lett.*, 122:180501, May 2019. DOI: 10.1103/PhysRevLett.122.180501. URL https://doi.org/10.1103/PhysRevLett.122.180501.

[23] Jarrod R McClean, Jonathan Romero, Ryan Babbush, and Alán Aspuru-Guzik. The theory of variational hybrid quantum-classical algorithms. *New Journal of Physics*, 18(2):023023, feb 2016. DOI: 10.1088/1367-2630/18/2/023023. URL https://doi.org/10.1088%2F1367-2630%2F18%2F2%2F023023.

[24] Jarrod R. McClean, Mollie E. Kimchi-Schwartz, Jonathan Carter, and Wibe A. de Jong. Hybrid quantum-classical hierarchy for mitigation of decoherence and determination of excited states. *Phys. Rev. A*, 95:042308, Apr 2017. DOI: 10.1103/PhysRevA.95.042308. URL https://doi.org/10.1103/PhysRevA.95.042308.

[25] Jarrod R McClean, Kevin J Sung, Ian D Kivlichan, Yudong Cao, Chengyu Dai, E Schuyler Fried, Craig Gidney, Brendan Gimby, Pranav Gokhale, et al. OpenFermion: the electronic structure package for quantum computers. *arXiv preprint arXiv:1710.07629*, 2017. URL https://arxiv.org/abs/1710.07629.

[26] Nikolaj Moll, Panagiotis Barkoutsos, Lev S Bishop, Jerry M Chow, Andrew Cross, Daniel J Egger, Stefan Filipp, Andreas Fuhrer, Jay M Gambetta, Marc Ganzhorn, et al. Quantum optimization using variational algorithms on near-term quantum devices. *Quantum Science and Technology*, 3(3):030503, jun 2018. DOI: 10.1088/2058-9565/aab822. URL https://doi.org/10.1088%2F2058-9565%2Faab822.

[27] 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, Jul 2016. DOI: 10.1103/PhysRevX.6.031007. URL https://doi.org/10.1103/PhysRevX.6.031007.

[28] Lawrence Page, Sergey Brin, Rajeev Motwani, and Terry Winograd. The PageRank citation ranking: Bringing order to the web. Technical Report 1999-66, Stanford InfoLab, November 1999. URL http://ilpubs.stanford.edu:8090/422/.

[29] Karl Pearson. On lines and planes of closest fit to systems of points in space. *The London, Edinburgh, and Dublin Philosophical Magazine and Journal of Science*, 2(11):559–572, 1901. DOI: 10.1080/14786440109462720. URL https://doi.org/10.1080/14786440109462720.

[30] Alberto Peruzzo, Jarrod McClean, Peter Shadbolt, Man-Hong Yung, Xiao-Qi Zhou, Peter J Love, Alán Aspuru-Guzik, and Jeremy L O'brien. A variational eigenvalue solver on a photonic quantum processor. *Nature communications*, 5, 2014. DOI: 10.1038/ncomms5213. URL https://doi.org/10.1038/ncomms5213.

[31] John Preskill. Quantum computing in the NISQ era and beyond. *Quantum*, 2:79, August 2018. ISSN 2521-327X. DOI: 10.22331/q-2018-08-06-79. URL https://doi.org/10.22331/q-2018-08-06-79.

[32] Jonathan Romero, Ryan Babbush, Jarrod R McClean, Cornelius Hempel, Peter J Love, and Alán Aspuru-Guzik. Strategies for quantum computing molecular energies using the unitary coupled cluster ansatz. *Quantum Science and Technology*, 4(1):014008, Oct 2018. DOI: 10.1088/2058-

9565/aad3e4. URL https://doi.org/10.1088%2F2058-9565%2Faad3e4.

[33] Nicholas C Rubin, Ryan Babbush, and Jarrod Mc-Clean. Application of fermionic marginal constraints to hybrid quantum algorithms. *New Journal of Physics*, 20(5):053020, may 2018. DOI: 10.1088/1367-2630/aab919. URL https://doi.org/10.1088%2F1367-2630%2Faab919.

[34] Ilya G Ryabinkin, Scott N Genin, and Artur F Izmaylov. Constrained variational quantum eigensolver: Quantum computer search engine in the Fock space. *Journal of chemical theory and computation*, 15(1):249–255, 2018. DOI: 10.1021/acs.jctc.8b00943. URL https://doi.org/10.1021/acs.jctc.8b00943.

[35] Raffaele Santagati, Jianwei Wang, Antonio A. Gentile, Stefano Paesani, Nathan Wiebe, Jarrod R. McClean, Sam Morley-Short, Peter J. Shadbolt, Damien Bonneau, Joshua W. Silverstone, David P. Tew, Xiaoqi Zhou, Jeremy L. O'Brien, and Mark G. Thompson. Witnessing eigenstates for quantum simulation of hamiltonian spectra. *Science Advances*, 4(1), 2018. DOI: 10.1126/sciadv.aap9646. URL https://doi.org/10.1126/sciadv.aap9646.

[36] Peter W. Shor. Polynomial-time algorithms for prime factorization and discrete logarithms on a quantum computer. *SIAM J. Comput.*, 26(5): 1484–1509, October 1997. ISSN 0097-5397. DOI: 10.1137/S0097539795293172. URL https://doi.org/10.1137/S0097539795293172.

[37] Damian S. Steiger, Thomas Häner, and Matthias Troyer. ProjectQ: an open source software framework for quantum computing. *Quantum*, 2: 49, January 2018. ISSN 2521-327X. DOI: 10.22331/q-2018-01-31-49. URL https://doi.org/10.22331/q-2018-01-31-49.

[38] Attila Szabo and Neil S. Ostlund. *Modern Quantum Chemistry: Introduction to Advanced Electronic Structure Theory*. Dover Publications, Inc., Mineola, first edition, 1996.

[39] Kristan Temme, Sergey Bravyi, and Jay M. Gambetta. Error mitigation for short-depth quantum circuits. *Phys. Rev. Lett.*, 119:180509, Nov 2017. DOI: 10.1103/PhysRevLett.119.180509. URL https://doi.org/10.1103/PhysRevLett.119.180509.

[40] Daochen Wang, Oscar Higgott, and Stephen Brierley. Accelerated variational quantum eigensolver. *Phys. Rev. Lett.*, 122:140504, Apr 2019. DOI: 10.1103/PhysRevLett.122.140504. URL https://doi.org/10.1103/PhysRevLett.122.140504.

[41] Dave Wecker, Matthew B. Hastings, and Matthias Troyer. Progress towards practical quantum vari-ational algorithms. *Phys. Rev. A*, 92:042303, Oct 2015. DOI: 10.1103/PhysRevA.92.042303. URL https://doi.org/10.1103/PhysRevA.92.042303.