Calculating energy derivatives for quantum chemistry on a quantum computer

O'Brien, T.E.; Senjean, B.; Sagastizabal, R.; Bonet Monroig, X.; Dutkiewicz, A.; Buda, F.; ...
; Visscher, L.

ARTICLE OPEN

Calculating energy derivatives for quantum chemistry on a quantum computer

Thomas E. O'Brien[1]*, Bruno Senjean[1,2], Ramiro Sagastizabal[3,4], Xavier Bonet-Monroig[1,3], Alicja Dutkiewicz[1,5], Francesco Buda[6], Leonardo DiCarlo[3,4] and Lucas Visscher [ID][2]

Modeling chemical reactions and complicated molecular systems has been proposed as the "killer application" of a future quantum computer. Accurate calculations of derivatives of molecular eigenenergies are essential toward this end, allowing for geometry optimization, transition state searches, predictions of the response to an applied electric or magnetic field, and molecular dynamics simulations. In this work, we survey methods to calculate energy derivatives, and present two new methods: one based on quantum phase estimation, the other on a low-order response approximation. We calculate asymptotic error bounds and approximate computational scalings for the methods presented. Implementing these methods, we perform geometry optimization on an experimental quantum processor, estimating the equilibrium bond length of the dihydrogen molecule to within 0.014 Å of the full configuration interaction value. Within the same experiment, we estimate the polarizability of the H_2 molecule, finding agreement at the equilibrium bond length to within 0.06 a.u. (2% relative error).

npj Quantum Information (2019)5:113 ; https://doi.org/10.1038/s41534-019-0213-4

INTRODUCTION

Quantum computers are at the verge of providing solutions for certain classes of problems that are intractable on a classical computer.[1] As this threshold nears, an important next step is to investigate how these new possibilities can be translated into useful algorithms for specific scientific domains. Quantum chemistry has been identified as a key area where quantum computers can stop being science and start doing science.[2–5] This observation has lead to an intense scientific effort towards developing and improving quantum algorithms for simulating time evolution[6,7] and calculating ground state energies[8–11] of molecular systems. Small prototypes of these algorithms have been implemented experimentally with much success.[10,12–15] However, advances over the last century in classical computational chemistry methods, such as density functional theory (DFT),[16] coupled cluster (CC) theory,[17] and quantum Monte-Carlo methods,[18] set a high bar for quantum computers to make impact in the field.

The ground and/or excited state energy is only one of the targets for quantum chemistry calculations. For many applications one also needs to be able to calculate the derivatives of the molecular electronic energy with respect to a change in the Hamiltonian.[19,20] For example, the energy gradient (or first-order derivative) for nuclear displacements is used to search for minima, transition states, and reaction paths[21] that characterize a molecular potential energy surface (PES). They also form the basis for molecular dynamics (MD) simulations to dynamically explore the phase space of the system in its electronic ground state[22] or, after a photochemical transition, in its electronically excited state.[23] While classical MD usually relies on force-fields which are parameterized on experimental data, there is a growing need to obtain these parameters on the basis of accurate quantum chemical calculations. One can easily foresee a powerful combination of highly accurate forces generated on a quantum computer with machine-learning algorithms for the generation of reliable and broadly applicable force-fields.[24] This route might be particularly important in exploring excited state PES and non-adiabatic coupling terms, which are relevant in describing light-induced chemical reactions.[25–27] Apart from these perturbations arising from changing the nuclear positions, it is also of interest to consider the effect that small external electric and/or magnetic fields have on the molecular energy. These determine well-known molecular properties, such as the (hyper)polarizability, magnetizability, A- and g-tensors, nuclear magnetic shieldings, among others.

Although quantum algorithms have been suggested to calculate derivatives of a function represented on a quantum register,[28–32] or of derivatives of a variational quantum eigensolver (VQE) for optimization purposes,[33,34] the extraction of molecular properties from quantum simulation has received relatively little focus. To the best of our knowledge only three investigations; in geometry optimization and molecular energy derivatives,[35] molecular vibrations,[36] and the linear response function;[37] have been performed to date.

In this work, we survey methods for the calculation of molecular energy derivatives on a quantum computer. We calculate estimation errors and asymptotic convergence rates of these methods, and detail the classical pre- and post-processing required to convert quantum computing output to the desired quantities. As part of this, we detail two new methods for such derivative calculations. The first involves simultaneous quantum phase and transition amplitude (or propagator) estimation, which we name "propagator and phase estimation" (PPE). The second is based on truncating the Hilbert space to an approximate (relevant) set of eigenstates, which we name the "eigenstate truncation approximation" (ETA). We use these methods to

[1]Instituut-Lorentz, Universiteit Leiden, P.O. Box 9506, 2300 RA Leiden, The Netherlands. [2]Division of Theoretical Chemistry, Vrije Universiteit Amsterdam, De Boelelaan 1083, 1081 HV Amsterdam, The Netherlands. [3]QuTech, Delft University of Technology, P.O. Box 5046, 2600 GA Delft, The Netherlands. [4]Kavli Institute of Nanoscience, Delft University of Technology, P.O. Box 5046, 2600 GA Delft, The Netherlands. [5]Faculty of Physics, University of Warsaw, ul. Pasteura 5, 02093 Warszawa, Poland. [6]Leiden Institute of Chemistry, Leiden University, Einsteinweg 55, P.O. Box 9502, 2300 RA Leiden, The Netherlands. *email: obrien@lorentz.leidenuniv.nl

Table 1. Calculated performance of the energy derivative estimation methods suggested in this work. The computation time scaling as a function of the estimation error ϵ is given, alongside an approximate range of scalings with respect to the system size N_{sys}. Calculation details, full approximation details, and intermediate steps are given in the Methods section.

	Hellmann–Feynman (first order)		PPE		ETA		Direct	
Time scaling with ϵ (fixed N_{sys})	ϵ^{-2}	†	ϵ^{-2}	*	ϵ^{-2}	†	ϵ^{-d+1}	†
	ϵ^{-2}	*			ϵ^{-2}	*	$\epsilon^{-\frac{d+2}{2}}$	*
	ϵ^{-1}	**			ϵ^{-1}	**	ϵ^{-1}	▽
Time scaling with N_{sys} (approx, fixed ϵ)	$N_{sys}^4 - N_{sys}^{13}$	†	$N_{sys}^4 - N_{sys}^{17.5}$	*	$N_{sys}^8 - N_{sys}^{21}$	†	$N_{sys}^4 - N_{sys}^{13}$	†
	$N_{sys}^4 - N_{sys}^{15}$	*			$N_{sys}^8 - N_{sys}^{23}$	*	$N_{sys}^2 - N_{sys}^7$	*
	$N_{sys}^{3.5} - N_{sys}^{11}$	**			$N_{sys}^{6.5} - N_{sys}^{17}$	**	$N_{sys}^2 - N_{sys}^7$	▽
Error sources unaccounted for	Basis set error		Basis set error		Basis set error		Basis set error	
	State error	(†, **)	Resolution error		Truncation error		State error	(†, **)
Req. $\hat{H}$ sim.	*, **		All		*, **		*, ▽	

†denotes performance estimates when using a VQE for state preparation, *denotes performance when using QPE for state preparation, **denotes performance if the amplitude amplification technique[39] is used (which requires a VQE for state preparation), and ▽denotes performing differentiation on a quantum computer using the methods of Kassal et al.[28,35] (which requires the ability to call the Hamiltonian as a quantum oracle as a function of the system parameters). Details of errors not accounted for are given in Methods. We further note whether methods require phase estimation (which requires long coherence times)

perform geometry optimization of the H_2 molecule on a super-conducting quantum processor, as well as its response to a small electric field (polarizability), and find excellent agreement with the full configuration interaction (FCI) solution.

RESULTS

Let $\hat{H}$ be a Hamiltonian on a $2^{N_{sys}}$-dimensional Hilbert space (e.g., the Fock space of an N_{sys}-spin orbital system), which has eigenstates

$$\hat{H}|\Psi_j\rangle = E_j|\Psi_j\rangle, \tag{1}$$

ordered by the energies E_j. In this definition, the Hamiltonian is parametrized by the specific basis set that is used and has additional coefficients $\lambda_1, \lambda_2, \ldots$, which reflect fixed external influences on the electronic energy (e.g., change in the structure of the molecule, or an applied magnetic or electric field). An dth-order derivative of the ground state energy with respect to the parameters λ_i is then defined as

$$D^{d_1, d_2, \ldots}_{\lambda_1, \lambda_2, \ldots} = \frac{\partial^d E_0(\lambda_1, \lambda_2, \ldots)}{\partial^{d_1}\lambda_1, \partial^{d_2}\lambda_2, \ldots}, \tag{2}$$

where $d = \sum_i d_i$. As quantum computers promise exponential advantages in calculating the ground state E_0 itself, it is a natural question to ask how to efficiently calculate such derivatives on a quantum computer.

The quantum chemical Hamiltonian

A major subfield of computational chemistry concerns solving the electronic structure problem. Here, the system takes a second-quantized ab initio Hamiltonian, written in a basis of molecular spinors $\{\phi_p(\mathbf{r})\}$ as follows

$$\hat{H} = \sum_{pq} h_{pq}\hat{E}_{pq} + \frac{1}{2}\sum_{pqrs} g_{pqrs}(\hat{E}_{pq}\hat{E}_{rs} - \delta_{q,r}\hat{E}_{ps}), \tag{3}$$

where $\hat{E}_{pq} = \hat{c}_p^\dagger \hat{c}_q$ and $\hat{c}_p^\dagger$ ($\hat{c}_p$) creates (annihilates) an electron in the molecular spinor ϕ_p. With Eq. (3) relativistic and nonrelativistic realizations of the method only differ in the definition of the matrix elements h_{pq} and g_{pqrs}.[38] A common technique is to assume pure spin-orbitals and integrate over the spin variable. As we want to develop a formalism that is also valid for relativistic calculations, we will remain working with spinors in this work. Adaptation to a

spinfree formalism is straightforward, and will not affect computational scaling and error estimates.

The electronic Hamiltonian defined above depends parametrically on the nuclear positions, both explicitly via the nuclear potential and implicitly via the molecular orbitals that change when the nuclei are displaced.

Asymptotic convergence of energy derivative estimation methods

In this section, we present and compare various methods for calculating energy derivatives on a quantum computer. In Table 1, we estimate the computational complexity of all studied methods in terms of the system size N_{sys} and the estimation error ϵ. We also indicate which methods require quantum phase estimation (QPE), as these require longer coherence times than variational methods. Many methods benefit from the amplitude estimation algorithm,[39] which we have included costings for. We approximate the scaling in N_{sys} between a best-case scenario (a lattice model with a low-weight energy derivative and aggressive truncation of any approximations), and a worst-case scenario (the electronic structure problem with a high-weight energy derivative and less aggressive truncation). The lower bounds obtained here are competitive with classical codes, suggesting that these methods will be tractable for use in large-scale quantum computing. However, the upper bounds will need reduction in future work to be practical, e.g., by implementing the strategies suggested in ref. [11,33,40]

For wavefunctions in which all parameters are variationally optimized, the Hellmann–Feynman theorem allows for ready calculation of energy gradients as the expectation value of the perturbing operator[35,41]

$$\frac{\partial E_0}{\partial \lambda} = \langle\Psi_0|\frac{\partial\hat{H}}{\partial\lambda}|\Psi_0\rangle. \tag{4}$$

This expectation value may be estimated by repeated measurement of a prepared ground state on a quantum computer (Section IV C), and classical calculation of the coefficients of the Hermitian operator $\partial\hat{H}/\partial\lambda$ (See Supplementary methods). If state preparation is performed using a VQE, estimates of the expectation values in Eq. (4) will often have already been obtained during the variational optimization routine. If one is preparing a state via QPE, one does not get these expectation values for free, and must repeatedly measure the quantum register on top of the phase

estimation routine. Such measurement is possible even with single-ancilla QPE methods which do not project the system register into the ground state (see Section IV I). Regardless of the state preparation method, the estimation error may be calculated by summing the sampling noise of all measured terms (assuming the basis set error and ground state approximation errors are minimal).

The Hellmann–Feynman theorem cannot be so simply extended to higher-order energy derivatives. We now study three possible methods for such calculations. The PPE method uses repeated rounds of QPE to measure the frequency-domain Green's function, building on previous work on Green's function techniques.[37,42,43] We may write an energy derivative via perturbation theory as a sum of products of path amplitudes A and energy coefficients f_A. For example, a second order energy derivative may be written as

$$\frac{\partial^2 E_0}{\partial\lambda_1\partial\lambda_2} = \langle\Psi_0|\frac{\partial^2\hat{H}}{\partial\lambda_1\partial\lambda_2}|\Psi_0\rangle$$
$$+ \sum_{j\neq 0} 2\,\mathrm{Re}\left[\langle\Psi_0|\frac{\partial\hat{H}}{\partial\lambda_1}|\Psi_j\rangle\langle\Psi_j|\frac{\partial\hat{H}}{\partial\lambda_2}|\Psi_0\rangle\right]\frac{1}{E_0 - E_j}, \quad (5)$$

allowing us to identify two amplitudes

$$\mathcal{A}_1(j) = \left\langle\Psi_0\left|\frac{\partial\hat{H}}{\partial\lambda_1}\right|\Psi_j\right\rangle\left\langle\Psi_j\left|\frac{\partial\hat{H}}{\partial\lambda_2}\right|\Psi_0\right\rangle, \quad (6)$$

$$\mathcal{A}_2 = \left\langle\Psi_0\left|\frac{\partial^2\hat{H}}{\partial\lambda_1\partial\lambda_2}\right|\Psi_0\right\rangle, \quad (7)$$

and two corresponding energy coefficients

$$f_1(E_0; E_j) = \frac{2}{E_0 - E_j}, f_2 = 1. \quad (8)$$

The generic form of a dth order energy derivative may be written as

$$D = \sum_{\mathcal{A}}\sum_{j_1}, \dots, j_{X_{\mathcal{A}}-1}\,\mathrm{Re}\left[\mathcal{A}(j_1, \dots, j_{X_{\mathcal{A}}-1})\right]$$
$$\times f_{\mathcal{A}}(E_0; E_{j_1}, \dots, E_{j_{X_{\mathcal{A}}-1}}), \quad (9)$$

where $X_{\mathcal{A}}$ counts the number of excitations in the path. As this is different from the number of responses of the wavefunction, $X_{\mathcal{A}}$ does not follow the $2n + 1$ rule; rather, $X_{\mathcal{A}} \leq d$. The amplitudes $\mathcal{A}$ take the form[1]

$$\mathcal{A}(j_1, \dots, j_{X_{\mathcal{A}}-1}) = \left\langle\Psi_0\left|\hat{P}_{X_{\mathcal{A}}}\prod_{x=1}^{X_{\mathcal{A}}-1}(|\Psi_{j_x}\rangle\langle\Psi_{j_x}|\hat{P}_x)\right|\Psi_0\right\rangle. \quad (10)$$

These may be estimated simultaneously with the corresponding energies E_{j_x} by applying rounds of QPE in between excitations by operators $\hat{P}_x$ (Section IV J). One may then classically calculate the energy coefficients f_A, and evaluate Eq. (9). Performing such calculation over an exponentially large number of eigenstates $|\Psi_{j_x}\rangle$ would be prohibitive. However, the quantum computer naturally bins small amplitudes of nearby energy with a resolution Δ controllable by the user. We expect the resolution error to be smaller than the error in estimating the amplitudes $\mathcal{A}(j_1, \dots, j_{X_{\mathcal{A}}-1})$ (Section IV L); we use the latter for the results in Table 1.

---

[1] Higher order ($d \geq 4$) amplitudes will eventually contain terms corresponding to disconnected excitations, which then are products of multiple terms of the form of Eq. (10). Our procedure may be extended to include these contributions; we have excluded them here for the sake of readability.

In lieu of the ability to perform the long circuits required for phase estimation, one may approximate the sum over (exponentially many) eigenstates $|\Psi_j\rangle$ in Eq. (9) by taking a truncated set of (polynomially many) approximate eigenstates $|\tilde{\Psi}_j\rangle$. We call such an approximation the eigenstate truncation approximation, or ETA for short. However, on a quantum computer, we expect both to better approximate the true ground state $|\Psi_0\rangle$, and to have a wider range of approximate excited states.[14,40,44–46] In this work, we focus on the quantum subspace expansion (QSE) method.[40] This method proceeds by generating a set of N_E vectors $|\chi_j\rangle$ connected to the ground state $|\Psi_0\rangle$ by excitation operators $\hat{E}_j$

$$|\chi_j\rangle = \hat{E}_j|\Psi_0\rangle. \quad (11)$$

This is similar to truncating the Hilbert space using a linear excitation operator in the (classical) equation of motion CC (EOMCC) approach.[47] The $|\chi_j\rangle$ states are not guaranteed to be orthonormal; the overlap matrix

$$S_{j,k}^{(\mathrm{QSE})} = \langle\chi_j|\chi_k\rangle, \quad (12)$$

is not necessarily the identity. To generate the set $|\tilde{\Psi}_j\rangle$ of orthonormal approximate eigenstates, one can calculate the projected Hamiltonian matrix

$$H_{j,k}^{(\mathrm{QSE})} = \langle\chi_j|\hat{H}|\chi_k\rangle, \quad (13)$$

and solve the generalized eigenvalue problem

$$\hat{H}^{(\mathrm{QSE})}\overrightarrow{v}^{(j)} = \tilde{E}_j\hat{S}^{(\mathrm{QSE})}\overrightarrow{v}^{(j)} \rightarrow |\tilde{\Psi}_j\rangle = \sum_l \overrightarrow{v}_l^{(j)}|\chi_l\rangle. \quad (14)$$

Regardless of the method used to generate the eigenstates $|\tilde{\Psi}_j\rangle$, the dominant computational cost of the ETA is the need to estimate N_E^2 matrix elements. Furthermore, to combine all matrix elements with constant error requires the variance of each estimation to scale as N_E^{-2} (assuming the error in each term is independent). This implies that, in the absence of amplitude amplification, the computational complexity scales as N_E^4. Taking all single-particle excitations sets $N_E \propto N_{\mathrm{sys}}^2$. However, in a lattice model one might consider taking only local excitations, setting $N_E \propto N_{\mathrm{sys}}$. Further reductions to N_E will increase the systematic error from Hilbert space truncation (Section IV M), although this may be circumvented somewhat by extrapolation.

For the sake of completeness, we also consider here the cost of numerically estimating an energy derivative by estimating the energy at multiple points

$$\frac{\partial^2 E}{\partial\lambda^2} = \frac{1}{\delta\lambda^2}(E(\lambda - \delta\lambda) + E(\lambda + \delta\lambda) - 2E(\lambda)) + O(\delta\lambda^2) \quad (15)$$

$$= \frac{1}{\delta\lambda}\left(\frac{\partial E}{\partial\lambda}(\lambda + \delta\lambda/2) - \frac{\partial E}{\partial\lambda}(\lambda - \delta\lambda/2)\right) + O(\delta\lambda^2). \quad (16)$$

Here, the latter formula is preferable if one has direct access to the derivative in a VQE via the Hellmann–Feynman theorem, while the former is preferable when one may estimate the energy directly via QPE. In either case, the sampling noise (Section IV C and Section IV G) is amplified by the division of $\delta\lambda$. This error then competes with the $O(\delta\lambda^2)$ finite difference error, the balancing of which leads to the scaling laws in Table 1. This competition can be negated by coherently copying the energies at different λ to a quantum register of L ancilla qubits and performing the finite difference calculation there.[28,48] Efficient circuits (and lower bounds) for the complexity of such an algorithm have not been determined, and proposed methods involve coherent calculation of the Hamiltonian coefficients on a quantum register. This would present a significant overhead on a near-term device, but with additive and better asymptotic scaling than the QPE step itself (which we use for the results in Table 1).
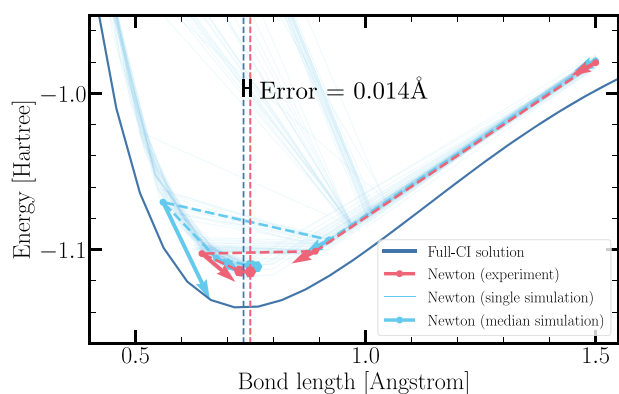
Fig. 1 Illustration of geometry optimization of the H_2 molecule. A classical optimization algorithm (Newton) minimizes the estimation of the true ground state energy (dark blue curve) on a superconducting transmon quantum computer (red crosses) as a function of the bond distance R_{H-H}. To improve convergence, the quantum computer provides estimates of the FCI gradient (red arrows) and the Hessian calculated with the response method. Dashed vertical lines show the position of the FCI and estimated minima (error 0.014Å). Light blue dashed lines show the median value of 100 density matrix simulations (Section IV O) of this optimization, with the shaded region the corresponding interquartile range.
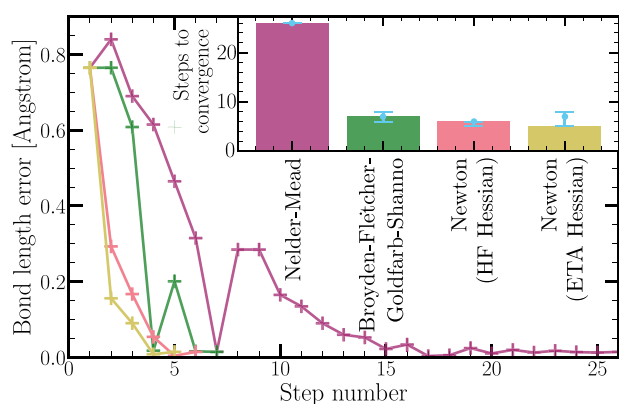
Fig. 2 Comparison of geometry optimization via different classical optimization routines, using a quantum computer to return energies and Jacobians as required, and estimating Hessians as required either via the ETA on the experimental device, or the Hartree–Fock (HF) approximation on a classical computer. Each algorithm was run till termination with a tolerance of 10^{-3}, so as to be comparable to the final error in the system. (Inset) bar plot of the number of function evaluations of the four compared methods. Light blue points correspond to median N_{fev} from 100 density-matrix simulations (Section IV O) of geometry optimization, and error bars to the interquartile ranges.

Geometry optimization on a superconducting quantum device

To demonstrate the use of energy derivatives directly calculated from a quantum computing experiment, we perform geometry optimization of the diatomic H_2 molecule, using two qubits of a superconducting transmon device. (Details of the experiment are given in Section IV N.) Geometry optimization aims to find the ground state molecular geometry by minimizing the ground state energy $E_0(\mathbf{R})$ as a function of the atomic co-ordinates R_i. In this small system, rotational and translational symmetries reduce this to a minimization as a function of the bond distance R_{H-H} In Fig. 1, we illustrate this process by sketching the path taken by Newton's minimization algorithm from a very distant initial bond distance ($R_{H-H} = 1.5$ Å). At each step of the minimization we show the
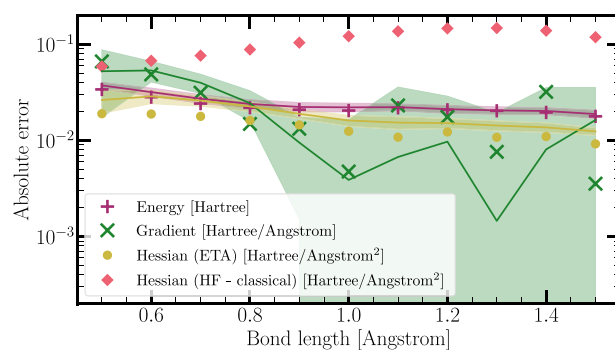

Fig. 3 Absolute error in energies and energy derivatives from an experimental quantum computation on 11 points of the bond dissociation curve of H_2. The error is dominated here by experimental sources (in particular qubit decay channels); error bars from sampling noise are smaller than the points themselves. Continuous lines connect the median value of 100 density matrix simulations at each points, with the shaded region corresponding to errors to the interquartile range.

gradient estimated via the Hellmann–Feynman theorem. Newton's method additionally requires access to the Hessian, which we calculated via the ETA (details given in Section IV N). The optimization routine takes 5 steps to converge to a minimum bond length of 0.749 Å, within 0.014 Å of the target FCI equilibrium bond length (given the chosen STO-3G basis set). To demonstrate the optimization stability, we performed 100 simulations of the geometry optimization experiment on the *quantumsim* density–matrix simulator,[49] with realistic sampling noise and coherence time fluctuations (details given in Section IV O). We plot all simulated optimization trajectories on Fig. 1, and highlight the median $(R_{H-H}, E(R_{H-H}))$ of the first 7 steps. Despite the rather dramatic variations between different gradient descent simulations, we observe all converging to within similar error bars, showing that our methods are indeed stable.

To study the advantage in geometry optimization from direct estimation of derivatives on a quantum computer, we compare in Fig. 2 our performance with gradient-free (Nelder–Mead) and Hessian-free (Broyden–Fletcher–Goldfarb–Shanno) optimization routines. We also compare the performance of the Newton's method with an approximate Hessian from Hartree–Fock (HF) theory. All methods converge to near-identical minima, but all methods using experimentally provided gradient information converge about four times as fast as Nelder–Mead. The density–matrix simulations predict that the ETA method Hessians provide less stable convergence than the HF Hessians; we attribute this to the fact that the HF Hessian at a fixed bond distance does not fluctuate between iterations. We also find that the separation between Hessian and Hessian-free gradient descent methods is insignificant in this one-dimensional problem. However we expect this to become more stark at larger system sizes, as is observed typically in numerical optimization.[50]

To separate the performance of the energy derivative estimation from the optimization routine, we study the error in the energy E, the Jacobian J and Hessian K given as $\epsilon_A = |A_{FCI} - A_{expt}|$, $(A = E, J, K)$. In Fig. 3, we plot these errors for different bond distances. For comparison we additionally plot the error in the HF Hessian approximation. We observe that the ETA Hessian is significantly closer than the HF-approximated Hessian to the true value, despite the similar performance in geometry optimization. The accuracy of the ETA improves at large bond distance, where the HF approximation begins to fail, giving hope that the ETA Hessian will remain appropriate in strongly correlated systems where this occurs as well. We also observe that the error in the ETA Hessian is approximately two times smaller than that of the energy, which we believe comes from error mitgation inherent in
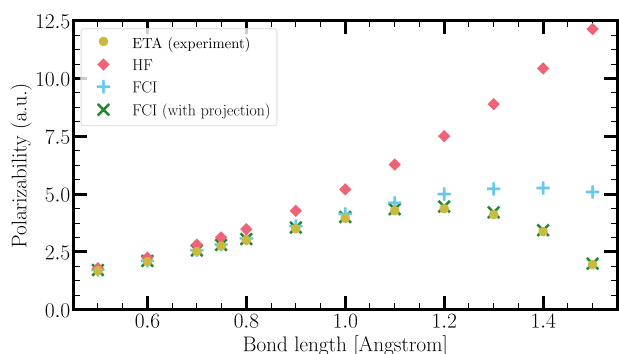
Fig. 4 Estimated polarizability of the hydrogen molecule as a function of the bond distance, in atomic units (1 a.u. = 0.14818471 Å³).

the QSE protocol.[40] The large fluctuations in the error in the gradient (large green shaded area) suggest that the estimation error in these points is mostly stochastic and not biased (unlike the error in the energy, which is variationally bound).

Polarizability estimation

A key property to model in quantum chemistry is the polarizability, which describes the tendency of an atom or molecule to acquire an induced dipole moment due to a change in an external electric field $\vec{F}$. The polarizability tensor may be calculated as $a_{i,j} = \frac{\partial E(\vec{F})}{\partial F_i \partial F_j}\Big|_{\vec{F}=0}$.[2] In Fig. 4, we calculate the z-component of the polarizability tensor of H_2 in the ETA, and compare it to FCI and HF polarizability calculations on a classical computer. We observe good agreement to the target FCI result at low R_{H-H}, finding a 0.060 a.u. (2.1%) error at the equilibrium bond distance (including the inaccuracy in estimating this distance). However, our predictions deviate from the exact result significantly at large bond distance ($R_{H-H} \gtrsim 1.2$ Å). We attribute this deviation to the transformation used to reduce the description of H_2 to a two-qubit device (see Section IV N), which is no longer valid when adding the dipole moment operator to the Hamiltonian. To confirm this, we classically compute the FCI polarizability following the same transformation (which corresponds to projecting the larger operator onto a 2-qubit Hilbert space). We find excellent agreement between this and the result from the quantum device across the entire bond dissociation curve. This implies that simulations of H_2 on a 4-qubit device should match the FCI result within experimental error.

DISCUSSION

In this work, we have surveyed possible methods for estimating energy gradients on a quantum computer, including two new techniques of our own design. We have estimated the computational complexity of these methods, both in terms of the accuracy required for the result and the size of the studied system. We have demonstrated the use of these methods on a small-scale quantum computing experiment, obtaining the equilibrium bond length of the H_2 molecule to 0.014 Å (2%) of the target FCI value, and estimating the polarizability at this bond length to within 0.060 a.u. (2.1%) of the same target.

Our methods do not particularly target the ground state over any other eigenstate of the system, and so can be used out-of-the-box for gradient estimation for excited state chemistry. They

hold further potential for calculating frequency-domain Green's functions in strongly correlated physics systems (as PPE estimates the gradient through a Green's function calculation). However, throughout this work we made the assumption that the gap δ between ground and excited state energies was sufficiently large to not be of concern (namely that $\delta \propto N_{sys}^{-1}$). Many systems of interest (such as high-temperature superconductors) are characterized by gap closings in the continuum limit. How this affects PPE is an interesting question for future work. Further investigation is also required to improve some of the results drawn upon for this work, in particular reducing the number of measurements required during a VQE and improving amplitude estimation during single-round QPE.

METHODS

Classical computation

The one- and two-electron integrals defining the fermionic Hamiltonian in Eq. (3) are obtained from a preliminary HF calculation that is assumed to be easily feasible on a classical computer. In non-relativistic theory the one-electron integrals are given by

$$h_{pq} = \int d\mathbf{r} \phi_p^*(\mathbf{r}) \left(-\frac{1}{2}\nabla_\mathbf{r} + V(\mathbf{r}) \right) \phi_q(\mathbf{r}), \tag{17}$$

where $V(\mathbf{r})$ is the electron-nuclear attraction potential from fixed nuclei at positions $\mathbf{R}_i$. The two-electron integrals are given by,

$$g_{pqrs} = \int\int d\mathbf{r}_1 d\mathbf{r}_2 \frac{\phi_p^*(\mathbf{r}_1)\phi_q(\mathbf{r}_1)\phi_r^*(\mathbf{r}_2)\phi_s(\mathbf{r}_2)}{r_{12}}. \tag{18}$$

For simplicity we used a finite difference technique to compute the matrix representations of perturbations corresponding to a change in nuclear coordinates and an external electric field

$$\frac{\partial \hat{H}}{\partial \lambda} \approx \frac{\hat{H}(\lambda + \delta\lambda/2) - \hat{H}(\lambda - \delta\lambda/2)}{\delta\lambda}, \tag{19}$$

and

$$\frac{\partial^2 \hat{H}}{\partial \lambda^2} \approx \frac{\hat{H}(\lambda + \delta\lambda) + \hat{H}(\lambda - \delta\lambda) - 2\hat{H}(\lambda)}{\delta\lambda^2}, \tag{20}$$

where $\delta\lambda = 0.001$ corresponds to a small change in λ. The above (perturbed) quantum chemical Hamiltonians have been determined within the Dirac program[51] and transformed into qubit Hamiltonians using the OpenFermion[52] package. This uses the newly developed, freely available[53] OpenFermion-Dirac interface, allowing for the simulation of relativistic quantum chemistry calculations on a quantum computer. While a finite difference technique was sufficient for the present purpose, such schemes are sensitive to numerical noise and have a high computational cost when applied to larger molecular systems. A consideration of the analytical calculation of energy derivatives can be found in the Supplementary methods.

Approximate bound calculation details

In this section we detail our method for calculating the approximate bounds in Table 1. We first estimate the error ϵ (Table 2, first row; details of the non-trivial calculations for the PPE and ETA methods given in Section IV H and Section IV M), respectively. Separately, we may calculate the time cost by multiplying the number of circuits, the number of repetitions of said circuits (n_m, N_m, and K depending on the method), and the time cost of each circuit (Table 2, second row). (This assumes access to only a single quantum processor, and can in some situations be improved by simultaneous measurement of commuting terms, as discussed in Section IV C.) We then choose the scaling of the number of circuit repetitions as a function of the other metaparameters to fix ϵ constant (Table 2, third row). We finally substitute the lower and upper bounds for these metaparameters in terms of the system size as stated throughout the remaining sections. For reference, we summarize these bounds in Table 3.

Quantum simulation of the electronic structure problem — preliminaries

To represent the electronic structure problem on a quantum computer, we need to rewrite the fermionic creation and annihilation operators $\hat{c}_i^\dagger$, $\hat{c}_i$ in

---

[2]The first-order derivative $\partial E/\partial F_i$ gives the dipole moment, which is also of interest, but is zero for the hydrogen molecule.

Table 2. Intermediate steps for the approximate calculation of the computational complexity as a function of the system size give in Table 1.

	Hellmann–Feynman (first order)		PPE		ETA		Direct	
Error scaling (ϵ)	$N_T^{\frac{1}{2}}n_m^{-\frac{1}{2}}$	†	$N_U^d \times$	*	$N_E N_T^{\frac{1}{2}} n_m^{-\frac{1}{2}}$	†	$(N_T n_m^{-1})^{\frac{1}{d+1}}$	†
	$N_U^{\frac{1}{2}}N_m^{-\frac{1}{2}}A_0^{-\frac{1}{2}}$	*	$\left(d^{\frac{d}{2}}\delta^{d-2}\Delta_2^{\frac{d}{2}}K^{-1}t^{-1}A_0^{-1}\right.$		$N_E N_U^{\frac{1}{2}}N_m^{-\frac{1}{2}}A_0^{-\frac{1}{2}}$	*	$(KA_0 t)^{\frac{-2}{d+2}}$	*
	$N_U^{\frac{1}{2}}K^{-1}$	**	$\left. + \delta^{d-1}\Delta^{\frac{d}{4}}N_m^{-\frac{1}{2}}A_0^{-\frac{1}{2}}\right)$		$N_E N_U^{\frac{1}{2}}K^{-1}$	**	$(KA_0 t)^{-1}$	∇
Time scaling	$N_T n_m T_P$	†	$dN_U^d KT_U$	*	$N_T N_E^2 n_m T_P$	†	$dN_T n_m T_P$	†
	$N_U KT_U$	*			$N_U N_E^2 KT_U$	*	dKT_U	*
	$N_U KT_P$	**			$N_U N_E^2 KT_P$	**	$2^d KT_U$	∇
Time scaling at fixed ϵ	$N_T^2 T_P$	†	$T_U \times$	*	$N_T^2 N_E^4 T_P$	†	$dN_T^2 T_P$	†
	$A_0^{-2}N_U^2 T_U$	*	$\left(d^{\frac{3}{2}}\delta^{d-2}t^{-1}\Delta_2^{\frac{d}{2}}A_0^{-1}N_U^{\frac{3d}{2}-1}\right.$		$N_U^2 N_E^4 A_0^{-2}T_U$	*	$dt^{-1}A_0^{-1}T_U$	*
	$N_U^{\frac{3}{2}}T_P$	**	$\left. + d\delta^{2d-2}\Delta^{d-1}A_0^{-\frac{1}{2}}N_U^{2d-1}\right)$		$N_U^{3/2}N_E^3 T_P$	**	$2^d t^{-1}A_0^{-1}T_U$	∇

Symbols are defined throughout the Methods section

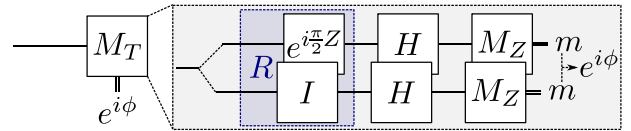**Table 3.** Summary of approximations used to derive the scaling laws with N sys in Table 1.

	Lower bound	Upper bound
T_P	N_{sys}^2	N_{sys}^5
T_U	1	N_{sys}^5
N_T	N_{sys}	N_{sys}^4
N_U	N_{sys}	N_{sys}^4
N_E	N_{sys}	N_{sys}^2
t^{-1}	N_{sys}	N_{sys}
A_0^{-1}	N_{sys}	N_{sys}
δ^{-1}	1	1
Δ^{-1}	1	1

Fig. 5 Definition of the circuit element M_T used throughout this work to estimate the phase $e^{i\phi}$ on a single qubit. This is done by repeatedly preparing and measuring the qubit along two different axis, by a combination of rotation and Hadamard gates and measurement M_Z in the computational basis. The final measurements may then be combined via Eq. (22).

terms of qubit operators (e.g., elements of the Pauli basis $\mathbb{P}^N = \{I, X, Y, Z\}^{\otimes N}$). This is necessarily a non-local mapping, as local fermionic operators anti-commute, while qubit operators commute. A variety of such transformations are known, including the Jordan–Wigner[54,55] and Bravyi–Kitaev[56] transformations, and more recent developments.[57–64]

After a suitable qubit representation has been found, we need to design quantum circuits to implement unitary transformations. Such circuits must be constructed from an appropriate set of primitive units, known as a (universal) gate-set. For example, one might choose the set of all single-qubit operators, and a two-qubit entangling gate such as the controlled-NOT, C-Phase, or iSWAP gates.[65] One can then build the unitary operators $e^{i\theta\hat{P}}$ for $\hat{P} \in \mathbb{P}^N$ exactly (in the absence of experimental noise) with a number of units and time linear in the size of $\hat{P}$.[66] (Here, size refers to the number of non-identity tensor factors of $\hat{P}$.) Optimizing the scheduling and size of these circuits is an open area of research, but many improvements are already known.[67]

Transformations of a quantum state must be unitary, which is an issue if one wishes to prepare, e.g., $\partial\hat{H}/\partial\lambda|\Psi_0\rangle$ on a quantum register ($\partial\hat{H}/\partial\lambda$ is almost always not unitary). To circumvent this, one must decompose $\partial\hat{H}/\partial\lambda$ as a sum of N_U unitary operators, perform a separate circuit for each unitary operator, and then combine the resulting measurements as appropriate. Such a decomposition may always be performed using the Pauli group (although better choices exist). Each such decomposition incurs a multiplicative cost of N_U to the computation time, and further increases the error in any final result by at worst a factor of $N_U^{1/2}$. This makes the computational complexities reported in Table 2 highly dependent on N_U. The scaling of N_U with the system size is highly dependent on the operator to be decomposed and the choice of decomposition. When approximating this in Table 3 we use a range

between $O(N_{sys})$ (which would suffice for a local potential in a lattice model) to $O(N_{sys}^4)$ (for a two-body interaction).

To interface with the outside world, a quantum register needs to be appropriately measured. Similarly to unitary transformations, one builds these measurements from primitive operations, typically the measurement of a single qubit in the Z basis. This may be performed by prior unitary rotation, or by decomposing an operator $\hat{O}$ into N_T Hermitian terms $\hat{O}_i$ (which may be measured separately). N_T differs from N_U defined above, as the first is for a Hermitian decomposition of a derivative operator and the second is for a unitary decomposition. Without a priori knowledge that the system is near an eigenstate of a operator $\hat{O}$ to be measured, one must perform n_m repeated measurements of each $\hat{O}_i$ to estimate $\langle\hat{O}\rangle$ to an accuracy $\propto n_m^{-1/2}N_T^{1/2}$. As such measurement is destructive, this process requires $n_m N_T$ preparations and pre-rotations on top of the measurement time. This makes the computational costs reported in Table 2 highly dependent on N_T. The scaling of N_T with the system size N_{sys} is highly dependent on the operator $\hat{O}$ to be measured and the choice of measurements to be made.[11,33] In Table 3, we assume a range between $O(N_{sys})$ and $O(N_{sys}^4)$ to calculate the approximate computation cost. This is a slight upper bound, as terms can be measured simultaneously if they commute, and error bounds may be tightened by accounting for the covariance between non-commuting terms.[11] The details on how this would improve the asymptotic scaling are still lacking in the literature, however, and so we leave this as an obvious target for future work.

Throughout this text we require the ability to measure a phase $e^{i\phi}$ between the $|0\rangle$ and $|1\rangle$ states of a single qubit. (This information is destroyed by a measurement in the Z basis, which may only obtain the amplitude on either state.) Let us generalize this to a mixed state on a single qubit, which has the density matrix[65]

$$\rho = \begin{pmatrix} p_0 & p_+ e^{i\phi} \\ p_+ e^{-i\phi} & p_1 \end{pmatrix}, \tag{21}$$

where $p_0 + p_1 = 1$, and $0 \leq p_+ \leq \sqrt{p_0 p_1} \leq 0.5$. If one repeatedly performs the two circuits in Fig. 5 (which differ by a gate $R = I$ or $R = R_Z = e^{i\frac{\pi}{4}Z}$), and

estimates the probability of a final measurement $m = 0, 1$, one may calculate

$$
\begin{aligned}
2p_+ e^{i\phi} &= P(m=0|R=I) - P(m=1|R=I) \\
&+ iP(m=0|R=R_Z) - iP(m=1|R=R_Z).
\end{aligned}
\tag{22}
$$

We define the circuit element M_T throughout this work as the combination of the two circuits to extract a phase using this equation. As above, the error in estimating the real and imaginary parts of $2p_+ e^{i\phi}$ scales as $n_m^{-1/2}$.

Hamiltonian Simulation

Optimal decompositions for complicated unitary operators are not in general known. For the electronic structure problem, one often wants to perform time evolution by a Hamiltonian $\hat{H}$, requiring a circuit for the unitary operator $U = e^{i\hat{H}t}$. For a local (fermionic or qubit) Hamiltonian, the length T_U of the required circuit is polynomial in the system size N_{sys}. However, the coefficient of this polynomial is often quite large; this depends on the chosen Hamiltonian, its basis set representation, the filling factor η (i.e., number of particles), and whether additional ancilla qubits are used.[4,5] Moreover, such circuits usually approximate the target unitary U with some $\tilde{U}$ with some bounds on the error $\epsilon_H = \| U - \tilde{U} \|_S$. This bound ϵ_H is proportional to the evolution time t, providing a 'speed limit' for such simulation.[68] For the electronic structure problem, current methods achieve scaling between $O(N_{sys}^2)$[69] and $O(N_{sys}^6)$[67,70] for the circuit length T_U, assuming $\eta \propto N_{sys}$ (and fixed t, ϵ). (When η is sublinear in N_{sys}, better results exist.[71]) The proven $O(N_{sys}^6)$ scaling is an upper bound, and most likely reduced by recent work.[72,73] For simpler models, such as the Hubbard model, scalings between $O(1)$ and $O(N_{sys})$ are available.[64,74] As we require $t \propto N_{sys}^{-1}$ for the purposes of phase estimation (described in Section IV G), this scaling is reduced by an additional factor throughout this work (though this cannot reduce the scaling below $O(1)$). For Table 3, we use a range of $T_U = O(1)$ and $T_U = O(N_{sys}^5)$ when approximating the scaling of our methods with the system size.

Ground state preparation and measurement

A key requirement for our derivative estimation methods is the ability to prepare the ground state $|\Psi_0\rangle$ or an approximation to it on the system register. Various methods exist for such preparation, including QPE (see Section IV G), adiabatic state preparation,[75] VQE,[10,11] and more recent developments.[76,77] Some of these preparation methods (in particular adiabatic and variational methods) are unitary, whilst others (phase estimation) are projective. Given a unitary preparation method, one may determine whether the system remains in the ground state by inverting the unitary and measuring in the computational basis (Section IV C). By contrast, such determination for QPE requires another phase estimation round, either via multiple ancilla qubits or by extending the methods in Section IV G. Unitary preparation methods have a slight advantage in estimating expectation values of unitary operators $\hat{U}$; the amplitude amplification algorithm[39] improves convergence of estimating $\langle \hat{U} \rangle$ from $\epsilon \propto T^{-1/2}$ to $\epsilon \propto T^{-1}$ (in a total computation time T). However, this algorithm requires repeated application of the unitary preparation while maintaining coherence, which is probably not achievable in the near-term. We list the computation time in Tables 1 and 2 both when amplitude amplification is (marked with $**$) and is not available.

Regardless of the method used, state preparation has a time cost that scales with the total system size. For QPE, this is the time cost KT_U of applying the required estimation circuits, where K is the total number of applications of $e^{i\hat{H}t}$.[78] The scaling of a VQE is dependent on the variational ansatz chosen.[11,33] The popular UCCSD ansatz for the electronic structure problem has a $O(N_{sys}^5)$ computational cost if implemented naively. However, recent work suggests aggressive truncation of the number of variational terms can reduce this as far as $O(N_{sys}^2)$.[33] We take this as the range of scalings for our approximations in Table 1.

Systematic error from ground state approximations (state error)

Traditionally, VQEs are not guaranteed to prepare the true ground state $|\Psi_0\rangle$, but instead some approximate ground state

$$
|\tilde{\Psi}_0\rangle = \sum_j a_j |\Psi_j\rangle.
\tag{23}
$$

Recent work has provided means of expanding VQEs iteratively to make the wavefunction error $1 - |a_0|^2$ arbitrarily small,[79–81] although it is unclear

how the time cost of these procedures scale with the system size. One may place very loose bounds on the error induces in the energy

$$
\begin{aligned}
2 \| \hat{H} \|_S (1 - |a_0|^2) &\geq |E_0 - \tilde{E}_0| = \sum_{j>0} a_j^* a_j (E_j - E_0) \\
&\geq \delta(1 - |a_0|^2) \geq 0,
\end{aligned}
\tag{24}
$$

where here $\| \hat{H} \|_S$ is the spectral norm of the Hamiltonian (its largest eigenvalue). (Note that while in general $\| \hat{H} \|_S$ is difficult to calculate, reasonable approximations are usually obtainable.) As $|\tilde{\Psi}_0\rangle$ is chosen to minimize the approximate energy $\tilde{E}_0$, one expects to be much closer to the smaller bound than the larger. For an operator $\hat{D}$ (such as a derivative operator $\partial \hat{H}/\partial \lambda$) other than the Hamiltonian, cross-terms will contribute to an additional error in the expectation value $D_0 = \langle |\hat{D}| \rangle$

$$
\begin{aligned}
|\tilde{D}_0 - D_0| = &\left| \sum_{i,j>0} a_i^* a_j \langle \Psi_i | \hat{D} | \Psi_j \rangle \right. \\
&\left. + 2 \operatorname{Re} \sum_j a_j^* a_0 \langle \Psi_j | \hat{D} | \Psi_0 \rangle + (|a_0|^2 - 1) D_0 \right|.
\end{aligned}
\tag{25}
$$

One can bound this above in turn using the fact that

$$
\sum_{i,j} a_i^* a_j \leq \left(\sum_i |a_i|^2 \right)^{1/2} \left(\sum_j |a_j|^2 \right)^{1/2} = (1 - |a_0|^2),
\tag{26}
$$

which leads to

$$
|\tilde{D}_0 - D_0| \leq 2 \| \hat{D} \|_S \left[(1 - |a_0|^2) + |a_0| \sqrt{1 - |a_0|^2} \right].
\tag{27}
$$

Combining this with the error in the energy gives the bound

$$
|\tilde{D}_0 - D_0| \leq 2 \| \hat{D} \|_S \left(\frac{|\tilde{E}_0 - E_0|^{1/2}}{\delta^{1/2}} + \frac{|\tilde{E}_0 - E_0|}{\delta} \right).
\tag{28}
$$

This ties the error in our derivative to the energy in our error, but with a square root factor that unfortunately slows down the convergence when the error is small. (This factor comes about precisely because we do not expect to be in an eigenstate of the derivative operator.) Unlike the above energy error, we cannot expect this bound to be loose without a good reason to believe that the orthogonal component $\sum_{j>0} a_j |\Psi_j\rangle$ has a similar energy gradient to the ground state. This will often not be the case; the low-energy excited state manifold is usually strongly coupled to the ground state by a physically-relevant excitation, causing the energies to move in opposite directions. Finding methods to circumvent this issue are obvious targets for future research. For example, one could optimize a VQE on a cost function other than the energy. One could also calculate the gradient in a reduced Hilbert space (see Section IV M) using eigenstates of $\hat{H}^{(QSE)} + \epsilon \hat{D}^{(QSE)}$ with small ϵ to ensure the coupling is respected.

Quantum phase estimation

Non-trivial measurement of a quantum computer is of similar difficulty to non-trivial unitary evolution. Beyond learning the expectation value of a given Hamiltonian $\hat{H}$, one often wishes to know specific eigenvalues E_i (in particular for the electronic structure problem, the ground and low-excited state energies). This may be achieved by QPE.[9] QPE entails repeated Hamiltonian simulation (as described above), conditional on an ancilla qubit prepared in the $|+\rangle = \frac{1}{\sqrt{2}}(|0\rangle + |1\rangle)$ state. (The resource cost in making the evolution conditional is constant in the system size.) Such evolution causes phase kickback on the ancilla qubit; if the system register was prepared in the state $\sum_j a_j |\Psi_j\rangle$, the combined (system plus ancilla) state evolves to

$$
\sum_j a_j |\Psi_j\rangle \otimes (|0\rangle + e^{ikE_j t} |1\rangle).
\tag{29}
$$

Repeated tomography at various k allows for the eigenphases E_j to be inferred, up to a phase $E_j t + 2\pi \equiv E_j t$. This inference can be performed directly with the use of multiple ancilla qubits,[9] or indirectly through classical post-processing of a single ancilla tomographed via the M_T gate of Fig. 5.[39,78,82–85]

The error in phase estimation comes from two sources; the error in Hamiltonian simulation and the error in the phase estimation itself. The error in Hamiltonian simulation may be bounded by ϵ_H (as calculated in the previous section), which in turn sets the time for a single unitary T_U.
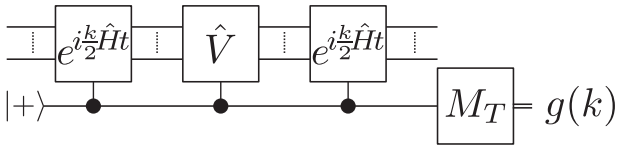
Fig. 6 A circuit to measure $\langle\Psi_j|\hat{V}|\Psi_j\rangle$ without preparing $|\Psi_j\rangle$ on the system register. The tomography box M_T is defined in Fig. 5.
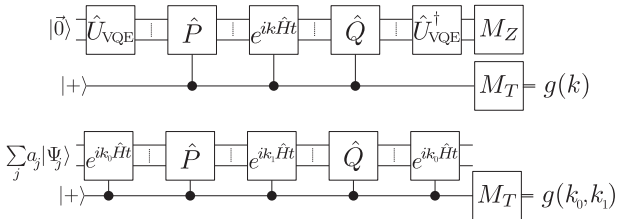
Fig. 7 Circuits for calculating path amplitudes (Eq. (10)) for a second-order derivative on a quantum computer (individual units described throughout Section IV C), using either a VQE (top) or QPE (bottom) for state preparation. Both circuits require an N-qubit system register and a single ancilla qubit. Repeat measurements of these circuit at different values of k (top) or k_0 and k_1 (bottom) allow for the inference of the amplitude, as described in the text. M_Z refers to a final measurement of all qubits in the computational basis, which is required for post-selection.

Assuming a sufficiently large gap to nearby eigenvalues, the optimal scaling of the error in estimating E_j is $A_j^{-1}t^{-1}K^{-1}$ (where $A_j = |a_j|^2$ is the amplitude of the jth eigenstate in the prepared state). Note that the phase equivalence $E_j t + 2\pi = E_j t$ sets an upper bound on t; in general we require $t \propto \| \hat{H} \|_s$, which typically scales with N_{sys}. (This scaling was incorporated into the estimates of T_U in the previous section.) The scaling of the ground state amplitude A_0 with the system size is relatively unknown, although numeric bounds suggest that it scales approximately as $1 - aN_{sys}^3$ for reasonable N_{sys}, with a a small constant. Approximating this as N_{sys}^{-1} implies that $K \propto N_{sys}^2$ is required to obtain a constant error in estimating the ground state energy.

The error in estimating an amplitude A_j during single-ancilla QPE has not been thoroughly investigated. A naive least-squares fit to dense estimation leads to a scaling of $n_m^{-1/2}k_{max}^{-1/3}$, where n_m is the number of experiments performed at each point $k = 1, \ldots, k_{max}$. One requires to perform controlled time evolution for up to $k_{max} \approx \max(t^{-1}, A_j^{-1})$ coherent steps in order to guarantee separation of ϕ_j from other phases. To obtain a constant error, one must then perform $n_m \propto k_{max}^{-\frac{2}{3}} \propto \max(A_j^{\frac{2}{3}}, t^{\frac{2}{3}})$ measurements at each k, implying $K = n_m k_{max} \propto \max(A_j^{-\frac{1}{3}}, t^{-\frac{1}{3}})$ applications of $e^{i\hat{H}t}$ are required. For the ground state ($A_j = A_0$), this gives scaling $K \propto N_{sys}^{\frac{1}{3}}$. By contrast, multiple-ancilla QPE requires N_m repetitions of $e^{i\hat{H}t}$ with $k_{max} = \max(A_0^{-1}, t^{-1})$ to estimate A_0 with an error of $(A_0(1 - A_0)N_m^{-1})^{1/2}$. This implies that $N_m \propto A_0$ measurements are sufficient, implying $K \propto N_m A_0^{-1}$ may be fixed constant as a function of the system size for constant error in estimation of A_0. Though this has not yet been demonstrated for single-round QPE, we expect it to be achievable and assume this scaling in this work.

The propagator and phase estimation method

In this section, we outline the circuits required and calculate the estimation error for our newly developed PPE method for derivative estimation. A prototype application of this method to a one-qubit system may be found in the Supplementary Notes of this work.

Estimating expectation values with single-ancilla QPE

Though single-ancilla QPE only weakly measures the system register, and does not project it into an eigenstate $|\Psi_j\rangle$ of the chosen Hamiltonian, it can still be used to learn properties of the eigenstates beyond their eigenvalues E_j. In particular, if one uses the same ancilla qubit to control

a further unitary operation $\hat{U}$ on the system register, the combined (system plus ancilla) state evolves from Eq. (29) to

$$\sum_{jj'} a_j(|0\rangle \otimes |\Psi_j\rangle + e^{ikE_j t}\langle\Psi_{j'}|\hat{U}|\Psi_j\rangle|1\rangle \otimes |\Psi_{j'}\rangle). \quad (30)$$

The phase accumulated on the ancilla qubit may then be calculated to be

$$g(k) = \sum_{jj'} a_j a_{j'}^* \langle\Psi_{j'}|\hat{U}|\Psi_j\rangle e^{ikE_j t}. \quad (31)$$

Note that the gauge degree of freedom is not present in Eq. (31); if one redefines $|\Psi_j\rangle \rightarrow e^{i\theta}|\Psi_j\rangle$, one must send $a_j \rightarrow e^{-i\phi_j}a_j$, and the phase cancels out. One may obtain $g(k)$ at multiple points k via tomography of the ancilla qubit (Fig. 5). From here, either Prony's method or Bayesian techniques may be used to extract phases $\omega_j \approx E_j t$ and corresponding amplitudes $a_j \approx \sum_{j'} a_j a_{j'}^* \langle\Psi_{j'}|\hat{U}|\Psi_j\rangle$.[85] The amplitudes a_j are often not terribly informative, but this changes if one extends this process over a family of operators U. For instance, if one chooses $U = e^{ik'\hat{H}t}\hat{V}e^{ik\hat{H}t}$ (with $\hat{V}$ unitary), an application of the Prony's method on k returns amplitudes of the form

$$a_j(k') \approx \sum_{j'} a_j a_{j'}^* e^{ik'E_{j'}t} \langle\Psi_{j'}|\hat{V}|\Psi_j\rangle, \quad (32)$$

from which a second application of Prony's method obtains phases $\omega_{j'} = E_{j'}t$ with corresponding amplitudes

$$a_{j,j'} \approx a_j a_{j'}^* \langle\Psi_{j'}|\hat{V}|\Psi_j\rangle. \quad (33)$$

Each subsequent application of QPE requires taking data with U^k fixed from $k = 1, \ldots, K$ to resolve K individual frequencies (and corresponding eigenvalues). However, if one is simply interested in expectation values $\langle\Psi_j|\hat{V}|\Psi_j\rangle$ (i.e., when $j = j'$), one may fix $k = k'$ and perform a single application of the Prony's method, reducing the number of circuits that need to be applied from $O(K^2)$ to $O(K)$ (see Fig. 6). The error in the estimator $a_{j,j}$ (Eq. (33)) may be bounded above by the error in the estimator a_j (Eq. (32)). However, to estimate $\langle\Psi_j|\hat{V}|\Psi_j\rangle$ from Eq. (33), one needs to divide by A_j. This propagates directly to the error, which then scales as $A_j^{-1/2}N_m^{-1/2}$. Thus constant error in estimating $\langle\Psi_j|\hat{V}|\Psi_j\rangle$ is achieved if $K \propto N_{sys}$.

PPE circuits

As presented, the operator $\hat{V}$ in Fig. 6 must be unitary. However if one applies additional phase estimation within $\hat{V}$ itself, one can extend this calculation to non-unitary operators, such as those given in Eq. (9). This is similar in nature to calculating the time-domain Green's function for small t on a quantum computer (which has been studied before in refs. [41–43]), but performing the transformation to frequency space with the Prony's method instead of a Fourier transform to obtain better spectral resolution. It can also be considered a generalization of ref. [37] beyond the linear response regime. To calculate a Xth order amplitude (Eq. (10)), one may set

$$\hat{V} = \left(\prod_{x=1}^{X-1} \hat{P}_x e^{ik_x\hat{H}t}\right)\hat{P}_X, \quad (34)$$

which is unitary if the $\hat{P}_x$ are chosen to be a unitary decomposition of $\partial\hat{H}/\partial\lambda_x$. In Fig. 7, we show two circuits for the estimation of a second order derivative with $\hat{P} = \partial\hat{H}/\partial\lambda_1$, $\hat{Q} = \partial\hat{H}/\partial\lambda_2$ (or some piece thereof). The circuits differ by whether QPE or a VQE is used for state preparation. If QPE is used for state preparation, the total phase accumulated by the ancilla qubit over the circuit is

$$g(k_0, k_1) = \sum_{j,m,n} a_m^* a_n \langle\Psi_m|\hat{P}|\Psi_j\rangle \langle\Psi_j|\hat{Q}|\Psi_n\rangle \times e^{ik_0 t(E_m + E_n)}e^{ik_1 tE_j}.$$

Applying the Prony's method at fixed k_1 will obtain a signal at phase $2tE_0$ with amplitude

$$a_{0,0}(k_1) \approx \sum_j a_0^* a_0 \langle\Psi_0|\hat{P}|\Psi_j\rangle \langle\Psi_j|\hat{Q}|\Psi_0\rangle e^{ik_1 tE_j}. \quad (35)$$

A second application of the Prony's method in k_1 allows us to obtain the required amplitudes

$$a_{0,j,0} \approx a_0^* a_0 \langle\Psi_0|\hat{P}|\Psi_j\rangle \langle\Psi_j|\hat{Q}|\Psi_0\rangle, \quad (36)$$

and the eigenvalues $\omega_j \approx E_j t$, allowing classical calculation of both the amplitudes and energy coefficients required to evaluate Eq. (9). If a VQE is used for state preparation instead, one must post-select on the system register being returned to $|\vec{0}\rangle$. Following this, the ancilla qubit will be in

the state

$$\frac{1}{\sqrt{2}}\left[|0\rangle + |1\rangle \sum_j e^{iktE_j}\langle\Psi_0^{(VQE)}|\hat{P}|\Psi_j\rangle\langle\Psi_j|\hat{Q}|\Psi_0^{(VQE)}\rangle\right],$$

with an accumulated phase $g(k) = a_{0,0}(k)$ (where $a_{0,0}$ is as defined above). Here, $|\Psi_0^{(VQE)}\rangle$ is the state prepared by the VQE unitary (which may not be the true ground state of the system). Both methods may be extended immediately to estimate higher-order amplitudes by inserting additional excitations and rounds of QPE, resulting in amplitudes of the form $a_{0,0}(k_1, \dots, k_X)$.

We note that the VQE post-selection does not constitute "throwing away data"; if the probability of post-selecting $|\Psi_0\rangle$ is p, we have

$$\sum_{k_1,\dots,k_X} |a_{0,0}(k_1, \dots, k_X)|^2 = p,$$

(37)

and as the variance in any term $a_{i,j}(k_1, \dots, k_X)$ scales as $|a_{i,j}(k_1, \dots, k_X)(1 - a_{i,j}(k_1, \dots, k_X))|$, the absolute error in estimating a derivative scales as $p^{1/2}$ (note the lack of minus sign). (Note here that the relative error scales as $p^{-1/2}$).

Energy discretization (resolution error)

The maximum number of frequencies estimatable from a signal $g(0), \dots, g(k)$ is $(k+1)/2$. (This can be seen by parameter counting; it differs from the bound of k for QPE[85] as the amplitudes are not real.) As the time required to obtain $g(k)$ scales at best linearly with k (Section IV D), we cannot expect fine enough resolution of all $2^{N_{sys}}$ eigenvalues present in a N_{sys}-qubit system. Instead, a small amplitude $\mathcal{A}(j_1, \dots, j_X)$ ($|\mathcal{A}(j_1, \dots, j_X)| \leq \Delta$) will be binned with paths $\mathcal{A}'(l_1, \dots, l_X)$ of similar energy ($\delta = \max_x |E_{j_x} - E_{l_x}| \ll \Delta$), and labeled with the same energy $E_{B_x} \approx E_{j_x} \approx E_{k_x}$.[85] Here, Δ is controlled by the length of the signal $g(k)$, i.e., $\Delta \propto \max(k)^{-1}$. This grouping does not affect the amplitudes; as evolution by $e^{ik\hat{H}t}$ does not mix eigenstates (regardless of energy), terms of the form $|\Psi_{j_x}\rangle\langle\Psi_{l_x}|$ do not appear. (This additional amplitude error would occur if one attempted to calculate single amplitudes of the form $\langle\Psi_j|\hat{P}|\Psi_k\rangle$ on a quantum device, e.g., using the method in Section IV G or that of ref. [37], and multiply them classically to obtain a $d > 1$st order derivative.) The PPE method then approximates Eq. (9) as

$$D = \sum_{\mathcal{A}} \sum_{B_1,\dots,B_{X_{\mathcal{A}}-1}} \mathcal{A}_{B_1,\dots,B_{X_{\mathcal{A}}-1}} f_{\mathcal{A}}(E_0; E_{B_1}, \dots, E_{B_{X_{\mathcal{A}}-1}}),$$

$$\mathcal{A}_{B_1,\dots,B_{X_{\mathcal{A}}-1}} = \sum_{j_x \in B_x} 2\,\text{Re}(\mathcal{A}(j_1, \dots, j_{X_{\mathcal{A}}-1})).$$

(38)

Classical post-processing then need only sum over the (polynomially many in Δ) bins B_x instead of the (exponentially many in N_{sys}) eigenstates $|\Psi_{j_x}\rangle$, which is then tractable.

To bound the resolution error in the approximation $f_{\mathcal{A}}(E_0; E_{j_1}, \dots, E_{j_{X_{\mathcal{A}}-1}}) \to f_{\mathcal{A}}(E_0; E_{B_1}, \dots, E_{B_{X_{\mathcal{A}}-1}})$, we consider the error if Format="TEX">{E}_{j} were drawn randomly from bins of width $\|\hat{H}\|_S \Delta$ (where $\|\hat{H}\|_S$ is the spectral norm). The energy functions f take the form of $X_{\mathcal{A}} - 1$ products of $\frac{1}{E_{j_x} - E_0}$. If each term is independent, these may be bounded as

$$\epsilon_f \leq X_{\mathcal{A}} \delta^{(X_{\mathcal{A}}-2)} \|\hat{H}\|_S \Delta,$$

(39)

where δ is the gap between the ground and excited states. Then, as the excitations $\hat{P}$ are unitary, for each amplitude $\mathcal{A}$ one may bound

$$\sum_{B_1,\dots,B_{X_{\mathcal{A}}-1}} |\mathcal{A}_{B_1,\dots,B_{X_{\mathcal{A}}}}|^2 \leq 1.$$

(40)

Propagating variances then obtains

$$\epsilon_D \leq d N_{\mathcal{A}}^{1/2} \delta^{d-2} \|\hat{H}\|_S \Delta,$$

(41)

where $N_{\mathcal{A}}$ is the number of amplitudes in the estimation of D. As we must decompose each operator into unitaries to implement in a circuit, $N_{\mathcal{A}} \propto N_U^d$.

This bound is quite loose; in general we expect excitations $\partial\hat{H}/\partial\lambda$ to couple to low-level excited states, which lie in a constant energy window (rather than one of width $\|\hat{H}\|_S$), and that contributions from different terms should be correlated (implying that $N_{\mathcal{A}}$ should be treated as

constant here). This implies that one may take Δ roughly constant in the system size, which we assume in this work.

Sampling noise error

We now consider the error in calculating Eq. (38) from a finite number of experiments (which is separate to the resolution error above). Following Section IV G we have that, if QPE is used for state preparation

$$\text{Var}[\mathcal{A}_{B_1,\dots,B_{X_{\mathcal{A}}-1}}] \propto |\mathcal{A}_{B_1,\dots,B_{X_{\mathcal{A}}-1}}| A_0^{-1} N_m^{-1}$$

(42)

$$\text{Var}[f(E_0; E_{B_1}, \dots, E_{B_{X_{\mathcal{A}}-1}})] \propto X_{\mathcal{A}} \delta^{2X_{\mathcal{A}}-4} K^{-2} t^{-2} |\mathcal{A}_{B_1,\dots,B_{X_{\mathcal{A}}-1}}|^{-2} A_0^{-2}.$$

(43)

If one were to use a VQE for state preparation, the factors of A_0 would be replaced by the state error of Section IV F. We have not included this calculation in Table 2 for the sake of simplicity. Then, assuming each term in Eq. (38) is independently estimated, we obtain

$$\text{Var}[D] = \sum_{\mathcal{A}} \sum_{B_1,\dots,B_{X_{\mathcal{A}}}} \left\{ \text{Var}[f_{\mathcal{A}}(E_0; E_{B_1}, \dots, E_{B_{X_{\mathcal{A}}}})] \left|\mathcal{A}_{B_1,\dots,B_{X_{\mathcal{A}}}}\right|^2 \right.$$

$$\left. + \text{Var}[\mathcal{A}_{B_1,\dots,B_{X_{\mathcal{A}}}}] \left|f_{\mathcal{A}}(E_0; E_{B_1}, E_{B_2}, \dots, E_{B_{X_{\mathcal{A}}}})\right|^2 \right\}.$$

(44)

Substituting the individual scaling laws one obtains

$$\text{Var}[D] \propto \sum_{\mathcal{A}} \sum_{B_1,\dots,B_{X_{\mathcal{A}}}} \left\{ X \delta^{2d-4} K^{-2} t^{-2} A_0^{-2} \right.$$

$$\left. + \delta^{2d-2} |\mathcal{A}_{B_1,\dots,B_{X_{\mathcal{A}}}}| A_0^{-1} N_m^{-1} \right\}$$

(45)

$$\leq N_{\mathcal{A}} \left\{ \delta^{2d-2} \Delta^{\frac{d}{2}} N_m^{-1} A_0^{-1} + d \delta^{2d-4} \Delta^d K^{-2} t^{-2} A_0^{-2} \right\},$$

(46)

where again $N_{\mathcal{A}} \propto N_U^d$. This result is reported in Table 2.

Eigenstate truncation approximation details

In this section, we outline the classical post-processing required to evaluate Eq. (9) in the ETA, using QSE to generate approximate eigenstates. We then calculate the complexity cost of such estimation, and discuss the systematic error in an arbitrary response approximation from Hilbert space truncation.

The chosen set of approximate excited states $|\tilde{\Psi}_j\rangle$ defines a subspace $\mathcal{H}^{(QSE)}$ of the larger FCI Hilbert space $\mathcal{H}^{(FCI)}$. To calculate expectation values within this subspace, we project the operators $\hat{O}$ of interest (such as derivatives like $\partial\hat{H}/\partial\lambda$) onto $\mathcal{H}^{(QSE)}$, giving a set of reduced operators $\hat{O}^{(QSE)}$ ($O_{i,j}^{(QSE)} = \langle\tilde{\Psi}_i|\hat{O}|\tilde{\Psi}_j\rangle$). These are $N_E \times N_E$-dimensional classical matrices, which may be stored and operated on in polynomial time. Methods to obtain the matrix elements $O_{i,j}^{(QSE)}$ are dependent on the form of the $|\tilde{\Psi}_j\rangle$ chosen. Within the QSE, one can obtain these by directly measuring[40]

$$\langle\chi_i|\hat{O}|\chi_j\rangle = \langle\Psi_0|\hat{E}_i^\dagger \hat{O}\hat{E}_j|\Psi_0\rangle,$$

(47)

using the techniques outlined in Section IV C, and rotating the basis from $\{|\chi_j\rangle\}$ to $\{|\tilde{\Psi}_j\rangle\}$ (using Eq. (14)).

The computational complexity for a derivative calculation within the QSE is roughly independent of the choice of $|\tilde{\Psi}_j\rangle$. The error ϵ may be bounded above by the error in each term of the $N_E \times N_E$ projected matrices, which scales as either $N_T^{1/2} n_m^{-1/2}$ (when directly estimating), $A_j^{-1/2} N_m^{-1/2}$ (when estimating via QPE), or $N_U^{1/2} K^{-1} n_m^{-1/2}$ (using the amplitude estimation algorithm). We assume that the N_E^2 terms are independently estimated, in which case ϵ scales with N_E. In general this will not be the case, and ϵ could scale as badly as N_E^2, but we do not expect this to be typical. Indeed, one can potentially use the covariance between different matrix elements to improve the error bound.[40] As we do not know the precise improvement this will provide, we leave any potential reduction in the computational complexity stated in Table 2 to future work. The calculation requires n_m repetitions of N_T circuits for each pair of N_E excitations, leading to a total number of $n_m N_T N_E^2$ preparations (each of which has a time cost T_P), as stated in Table 2. (With the amplitude amplification algorithm, the dominant time cost comes from running $O(N_T N_E^2)$ circuits of length KT_P.)

Regardless of the method of generating eigenstates, the ETA incurs a systematic truncation error from approximating an exponentially large number of true eigenstates $|\Psi_j\rangle$ by a polynomial number of approximate

eigenstates $|\tilde{\Psi}_j\rangle = \sum_l \tilde{A}_{j,l}|\Psi_l\rangle$. This truncation error can be split into three pieces. Firstly, an excitation $\hat{P}|\Psi_0\rangle$ may not lie within the response subspace $\mathcal{H}^{(QSE)}$, in which case the pieces lying outside the space will be truncated away. Secondly, the term $\hat{P}|\tilde{\Psi}_j\rangle\langle\tilde{\Psi}_j|\hat{Q}$ may contain terms of the form $\hat{P}|\Psi_j\rangle\langle\Psi_l|\hat{Q}$, which do not appear in the original resolution of the identity. Thirdly, the approximate energies $\tilde{E}_j$ may not be close to the true energies E_j (especially when $|\tilde{\Psi}_j\rangle$ is a sum of true eigenstates $|\Psi_l\rangle$ with large energy separation $E_j - E_l$). If one chooses excitation operators $\hat{E}_j$ in the QSE so that $\hat{P} = \sum_j p_j \hat{E}_j$, one completely avoids the first error source. By contrast, if one chooses a truncated set of true eigenstates $|\tilde{\Psi}_j\rangle = |\Psi_j\rangle$, one avoids the second and third error sources exactly. In the Supplementary Notes of this work, we expand on this point, and place some loose bounds on these error sources.

Experimental methods

The experimental implementation of the geometry optimization algorithm was performed using two of three transmon qubits in a circuit QED quantum processor. This is the same device used in ref. [86] (raw data is the same as in Fig. 1e of this paper, but with heavy subsequent processing). The two qubits have individual microwave lines for single-qubit gating and flux-bias lines for frequency control, and dedicated readout resonators with a common feedline. Individual qubits are addressed in readout via frequency multiplexing. The two qubits are connected via a common bus resonator that is used to achieve an exchange gate

$$\begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & \cos(\theta) & i\sin(\theta) & 0 \\ 0 & i\sin(\theta) & \cos(\theta) & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}, \tag{48}$$

via a flux pulse on the high-frequency qubit, with an uncontrolled additional single-qubit phase that was canceled out in post-processing. The exchange angle θ may be fixed to a $\pi/6000$ resolution by using the pulse duration (with a 1 ns duration) as a rough knob and fine-tuning with the pulse amplitude. Repeat preparation and measurement of the state generated by exciting to $|01\rangle$ and exchanging through one of 41 different choices of θ resulted in the estimation of 41 two-qubit density matrices ρ_i via linear inversion tomography of 10^4 single-shot measurements per prerotation.[87] All circuits were executed in eQASM[88] code compiled with the QuTech OpenQL compiler, with measurements performed using the qCoDeS[89] and PycQED[90] packages.

To use the experimental data to perform geometry optimization for H_2, the ground state was estimated via a VQE.[10,11] The Hamiltonian at a given H–H bond length R_{H-H} was calculated in the STO-3G basis using the Dirac package,[51] and converted to a qubit representation using the Bravyi–Kitaev transformation, and reduced to two qubits via exact block-diagonalization[12] using the Openfermion package[52] and the Openfermion–Dirac interface.[53] With the Hamiltonian $\hat{H}(R_{H-H})$ fixed, the ground state was chosen variationally: $\rho(R_{H-H}) = \min_{\rho_i} \text{Trace}[\hat{H}(R_{H-H})\rho_i]$. The gradient and Hessian were then calculated from $\rho(R_{H-H})$ using the Hellmann–Feynman theorem (Section II B) and ETA (Section IV M), respectively. For the ETA, we generated eigenstates using the QSE, with the Pauli operator XY as a single excitation. This acts within the number conserving subspace of the two-qubit Hilbert space, and, being imaginary, will not have the real-valued H_2 ground state as an eigenstate. (This in turn guarantees the generated excited state is linearly independent of the ground state.) For future work, one would want to optimize the choice of θ at each distance R_{H-H}, however this was not performed due to time constraints. We have also not implemented the error mitigation strategies studied in ref. [86] for the sake of simplicity.

Simulation methods

Classical simulations of the quantum device were performed in the full-density–matrix simulator (quantumsim).[49] A realistic error model of the device was built using experimentally calibrated parameters to account for qubit decay (T_1), pure dephasing (T_2^*), residual excitations of both qubits, and additional dephasing of qubits fluxed away from the sweet spot (which reduces T_2^* to $T_2^{*,red}$ for the duration of the flux pulse). This error model further accounted for differences in the observed noise model on the individual qubits, as well as fluctuations in coherence

times and residual excitation numbers. Further details of the error model may be found in ref. [86] (with device parameters in Table S1 of this reference).

With the error model given, 100 simulated experiments were performed at each of the 41 experimental angles given. Each experiment used unique coherence time and residual excitation values (drawn from a distribution of the observed experimental fluctuations), and had appropriate levels of sampling noise added. These density matrices were then resampled 100 times for each simulation.

DATA AVAILABILITY

Experimental data and code for post-processing are available from the corresponding author upon reasonable request.

CODE AVAILABILITY

All packages used in this work are open source and available via git repositories online.

REFERENCES

1. Preskill, J. Quantum computing in the NISQ era and beyond. *Quantum* **2**, 79 (2018).
2. Lloyd, S. Universal quantum simulators. *Science* **273**, 1073–1078 (1996).
3. Reiher, M., Wiebe, N., Svore, K. M. & Wecker, D. Elucidating reaction mechanisms on quantum computers. *Proc. Natl Acad. Sci. USA* **114**, 7555–7560 (2017).
4. Cao, Y. et al. Quantum chemistry in the age of quantum computing. *Chem. Rev.* **119**, 10856–10915 (2019).
5. McArdle, S. et al. *Quantum Computational Chemistry*. Preprint at: arXiv:1808.10402. https://arxiv.org/abs/1808.10402 (2018).
6. Abrams, D. S. & Lloyd, S. Simulation of many-body fermi systems on a universal quantum computer. *Phys. Rev. Lett.* **79**, 2586 (1997).
7. Zalka, C. Simulating quantum systems on a quantum computer. *Proc. Royal Soc. Lond. A* **454**, 313–322 (1998).
8. Aspuru-Guzik, A., Dutoi, A. D., Love, P. J. & Head-Gordon, M. Simulated quantum computation of molecular energies. *Science* **309**, 1704–1707 (2005).
9. Kitaev, A. Y. *Quantum Measurements and the Abelian Stabilizer Problem*. Preprint at: arXiv:quant-ph/9511026. https://arxiv.org/abs/quant-ph/9511026 (1995).
10. Peruzzo, A. et al. A variational eigenvalue solver on a photonic quantum processor. *Nat. Commun.* **5**, 4213 (2014).
11. McClean, J. R., Romero, J., Babbush, R. & Aspuru-Guzik, A. The theory of variational hybrid quantum-classical algorithms. *New J. Phys.* **18**, 023023 (2016).
12. O'Malley, P. J. J. et al. Scalable quantum simulation of molecular energies. *Phys. Rev. X* **6**, 031007 (2016).
13. Kandala, A. et al. Hardware-efficient variational quantum eigensolver for small molecules and quantum magnets. *Nature* **549**, 242 (2017).
14. Santagati, R. et al. Witnessing eigenstates for quantum simulation of hamiltonian spectra. *Sci. Adv.* **4**, eaap9646 (2018).
15. Colless, J. I. et al. Computation of molecular spectra on a quantum processor with an error-resilient algorithm. *Phys. Rev. X* **8**, 011021 (2018).
16. Dreizler, R. M. & Gross, E. K. U. *Density Functional Theory: An Approach to the Quantum Many-Body Problem* (Springer: Berlin. Heidelberg, 1990).
17. Shavitt, I. & Bartlett, R. J. *Many-Body Methods in Chemistry and Physics: MBPT and Coupled-Cluster Theory* (Cambridge University Press, 2009).
18. Booth, G. H., Thom, A. J. & Alavi, A. Fermion monte carlo without fixed nodes: a game of life, death, and annihilation in slater determinant space. *J. Chem. Phys.* **131**, 054106 (2009).
19. Jensen, F. *Introduction to Computational Chemistry*, 2nd ed. (John Wiley & Sons, 2007).
20. Norman, P., Ruud, K. & Saue, T. *Principles and Practices of Molecular Properties: Theory, Modeling, and Simulations*. https://doi.org/10.1002/9781118794821 (John Wiley & Sons, 2018).
21. Schlegel, H. B. Geometry optimization. *Wiley Interdiscip. Rev. Comput. Mol. Sci.* **1**, 790–809 (2011).
22. Marx, D. & Hutter, J. *Ab Initio Molecular Dynamics: Basic Theory and Advance Methods*. https://www.cambridge.org/nl/academic/subjects/physics/mathematical-methods/ab-initio-molecular-dynamics-basic-theory-and-advanced-methods?format=PB (Cambridge University Press, 2009).
23. Tully, J. C. Molecular dynamics with electronic transitions. *J. Chem. Phys.* **93**, 1061–1071 (1990).

24. Behler, J. Perspective: machine learning potentials for atomistic simulations. *J. Chem. Phys.* **145**, 170901 (2016).

25. Li, Z. & Liu, W. First-order nonadiabatic coupling matrix elements between excited states: a lagrangian formulation at the cis, rpa, td-hf, and td-dft levels. *J. Chem. Phys.* **141**, 014110 (2014).

26. Curchod, B. F., Rothlisberger, U. & Tavernelli, I. Trajectory-based nonadiabatic dynamics with time-dependent density functional theory. *ChemPhysChem* **14**, 1314–1340 (2013).

27. Faraji, S., Matsika, S. & Krylov, A. I. Calculations of non-adiabatic couplings within equation-of-motion coupled-cluster framework: theory, implementation, and validation against multi-reference methods. *J. Chem. Phys.* **148**, 044103 (2018).

28. Jordan, S. P. Fast quantum algorithm for numerical gradient estimation. *Phys. Rev. Lett.* **95**, 050501 (2005).

29. Gilyén, A., Arunachalam, S. & Wiebe, N. Optimizing quantum optimization algorithms via faster quantum gradient computation. *Proc. Symp. Disc. Alg.* **30**, 1425–1444 (2019).

30. Dallaire-Demers, P.-L., Romero, J., Veis, L., Sim, S. & Aspuru-Guzik, A. Low-depth circuit ansatz for preparing correlated fermionic states on a quantum computer. *Quant. Sci. Tech* **4**, 045005 (2019).

31. Schuld, M., Bergholm, V., Gogolin, C., Izaac, J. & Killoran, N. Evaluating analytic gradients on quantum hardware. *Phys. Rev. A* **99**, 032331 (2019).

32. Harrow, A. & Napp, J. *Low-Depth Gradient Measurements Can Improve Convergence in Variational Hybrid Quantum-Classical Algorithms*. Preprint at arXiv:1901.05374. https://arxiv.org/abs/1901.05374 (2019).

33. Romero, J. et al. Strategies for quantum computing molecular energies using the unitary coupled cluster ansatz. *Quantum Sci. Technol.* **4**, 014008 (2018).

34. Guerreschi, G. G. & Smelyanskiy, M. *Practical Optimization for Hybrid Quantum-Classical Algorithms*. Preprint at: arXiv:1701.01450. https://arxiv.org/abs/1701.01450 (2017).

35. Kassal, I. & Aspuru-Guzik, A. Quantum algorithm for molecular properties and geometry optimization. *J. Chem. Phys.* **131**, 224102 (2009).

36. McArdle, S. Mayorov, A. Shan, X. Benjamin, S. & Yuan, X. Quantum Computation of Molecular Vibrations. *Chem. Sci.* **10**, 5725–5735 (2019).

37. Roggero, A. Carlson, J. Linear Response on a Quantum Computer. *Phys. Rev. C* **100**, 034610 (2019).

38. Visscher, L. The Dirac equation in quantum chemistry: strategies to overcome the current computational problems. *J. Comp. Chem.* **23**, 759–766 (2002).

39. Knill, E., Ortiz, G. & Somma, R. D. Optimal quantum measurements of expectation values of observables. *Phys. Rev. A* **75**, 012328 (2007).

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

41. Wecker, D. et al. Solving strongly correlated electron models on a quantum computer. *Phys. Rev. A* **92**, 062318 (2015).

42. Dallaire-Demers, P.-L. & Wilhelm, F. K. Method to efficiently simulate the thermodynamic properties of the fermi-hubbard model on a quantum computer. *Phys. Rev. A* **93**, 032303 (2016).

43. Bauer, B., Wecker, D., Millis, A. J., Hastings, M. B. & Troyer, M. Hybrid quantum-classical approach to correlated materials. *Phys. Rev. X* **6**, 031045 (2016).

44. Parrish, R. M., Hohenstein, E. G., McMahon, P. L. & Martinez, T. J. Quantum computation of electronic transitions using a variational quantum Eigensolver. *Phys. Rev. Lett.* **122**, 230401 (2019).

45. Higgott, O., Wang, D. & Brierley, S. Variational quantum computation of excited states. *Quantum* **3**, 156 (2019).

46. Endo, S., Jones, T., McArdle, S., Yuan, X. & Benjamin, S. Variational quantum algorithms for discovering Hamiltonian spectra. *Phys. Rev. A* **99**, 062304 (2019)

47. Bartlett, R. J. Coupled cluster theory and its equation of motion extensions. *Wiley Interdiscip. Rev.* **2**, 126–211 (2011).

48. Kassal, I., Jordan, S. P., Love, P. J., Mohseni, M. & Aspuru-Guzik, A. Polynomial-time quantum algorithm for the simulation of chemical dynamics. *Proc. Natl Acad. Sci. USA* **105**, 18681 (2008).

49. O'Brien, T., Tarasinski, B. & DiCarlo, L. Density-matrix simulation of small surface codes under current and projected experimental noise. *npj Quant. Inf.* **3**. https://doi.org/10.1038/s41534-017-0039-x (2017).

50. Nocedal, J. & Wright, S. *Numerical Optimization* (Springer Science & Business Media, 2006).

51. Saue, T. et al. DIRAC, a relativistic ab initio electronic structure program, Release DIRAC18. Available at: https://doi.org/10.5281/zenodo.2253986, see also http://www.diracprogram.org (2018).

52. McClean, J. R. et al. Openfermion: the electronic structure package for quantum computers. Preprint at: arXiv:1710.07629. https://arxiv.org/abs/1710.07629 (2017).

53. Senjean, B. https://github.com/bsenjean/Openfermion-Dirac.

54. Jordan, P. & Wigner, E. Über das Paulische Äquivalenzverbot. *Z. Phys.* **47**, 631 (1928).

55. Ortiz, G., Gubernatis, J., Knill, E. & Laflamme, R. Quantum algorithms for fermionic simulations. *Phys. Rev. A* **64**, 022319 (2001).

56. Bravyi, S. B. & Kitaev, A. Y. Fermionic quantum computation. *Ann. Phys.* **298**, 210–226 (2002).

57. Verstraete, F. & Cirac, J. I. Mapping local Hamiltonians of fermions to local Hamiltonians of spins. *J. Stat. Mech. Theory Exp.* **2005**, P09012 (2005).

58. Seeley, J. T., Richard, M. J. & Love, P. J. The Bravyi-Kitaev transformation for quantum computation of electronic structure. *J. Chem. Phys.* **137**, 224109 (2012).

59. Whitfield, J. D., Havlíček, V. & Troyer, M. Local spin operators for fermion simulations. *Phys. Rev. A* **94**, 030301 (2016).

60. Steudtner, M. & Wehner, S. Fermion-to-qubit mappings with varying resource requirements for quantum simulation. *New J. Phys.* **20**, 063010 (2018).

61. Setia, K., Bravyi, S., Mezzacapo, A. & Whitfield, J. D. Superfast encodings for Fermionic quantum simulation. *Phys. Rev. Res.* **1**, 033033 (2019).

62. Setia, K. & Whitfield, J. D. Bravyi-Kitaev superfast simulation of electronic structure on a quantum computer. *J. Chem. Phys.* **148**, 164104 (2018).

63. Tranter, A., Love, P. J., Mintert, F. & Coveney, P. V. A comparison of the Bravyi-Kitaev and Jordan-Wigner transformations for the quantum simulation of quantum chemistry. *J. Chem. Theory Comput.* **14**, 5617–5630 (2018).

64. O'Brien, T., Rożek, P. & Akhmerov, A. Majorana-based fermionic quantum computation. *Phys. Rev. Lett.* **120**, 220504 (2018).

65. Nielsen, M. A. & Chuang, I. L. *Quantum Computation and Quantum Information* (Cambridge University Press, 2002). https://doi.org/10.1017/CBO9780511976667.

66. Whitfield, J. D., Biamonte, J. & Aspuru-Guzik, A. Simulation of electronic structure hamiltonians using quantum computers. *Mol. Phys.* **109**, 735–750 (2011).

67. Hastings, M. B., Wecker, D., Bauer, B. & Troyer, M. Improving quantum algorithms for quantum chemistry. *Quant. Inf. Comp.* **15**, 1 (2015).

68. Berry, D., Ahokas, G., Cleve, R. & Sanders, B. Efficient quantum algorithms for simulating sparse hamiltonians. *Comm. Math. Phys.* **270**, 359–371 (2007).

69. Low, G. & Wiebe, N. *Hamiltonian Simulation in the Interaction Picture*. Preprint at: arXiv:1805.00675. https://arxiv.org/abs/1805.00675 (2018).

70. Motzoi, F., Kaicher, M. & Wilhelm, F. Linear and logarithmic time compositions of quantum many-body operators. *Phys. Rev. Lett.* **119**, 160503 (2017).

71. Babbush, R., Berry, D., McClean, J. & Neven, H. Quantum simulation of chemistry with sublinear scaling to the continuum. *npj Quant. Inf.* **5**, 92 (2019).

72. Motta, M. et al. *Low Rank Representations for Quantum Simulation of Electronic Structure*. Preprint at: arXiv:1808.02625. https://arxiv.org/abs/1808.02625 (2018).

73. Campbell, E. A Random Compiler for Fast Hamiltonian Simulation. *Phys. Rev. Lett.* **123**, 070503 (2019).

74. Kivlichan, I. et al. Improved fault-tolerant quantum simulation of condensed-phase correlated electrons via trotterization. Preprint at: arXiv:1902.10673. https://arxiv.org/abs/1902.10673 (2019).

75. Wu, L.-A. Byrd, M. & Lidar, D. Polynomial-time simulation of pairing models on a quantum computer. *Phys. Rev. Lett.* **89**. https://journals.aps.org/prl/abstract/10.1103/PhysRevLett.89.057904 (2002).

76. Motta, M. et al. *Quantum Imaginary Time Evolution, Quantum Lanczos, and Quantum Thermal Averaging*. Preprint at: arXiv:1901.07653. https://arxiv.org/abs/1901.07653 (2019).

77. Kyriienko, O. Quantum inverse iteration algorithm for near-term quantum devices. Preprint at: arXiv:1901.09988. https://arxiv.org/abs/1901.09988 (2019).

78. Higgins, B. L. et al. Demonstrating Heisenberg-limited unambiguous phase estimation without adaptive measurements. *New J. Phys.* **11**, 073023 (2009).

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

80. Ryabinkin, I. G. & Genin, S. N. *Iterative Qubit Coupled Cluster Method: A Systematic Approach to the Full-ci Limit in Quantum Chemistry Calculations on Nisq Devices*. Preprint at: arXiv:1906.11192. https://arxiv.org/abs/1906.11192 (2019).

81. Herasymenko, Y. & O'Brien, T.E. *A Diagrammatic Approach to Variational Quantum Ansatz Construction*. Preprint at: arXiv:1907.08157. https://arxiv.org/abs/1907.08157 (2019).

82. Svore, K., Hastings, M. & Freedman, M. Faster phase estimation. *Quant. Inf. Comp.* **14**, 306–328 (2013).

83. Kimmel, S., Low, G. H. & Yoder, T. J. Robust calibration of a universal single-qubit gate set via robust phase estimation. *Phys. Rev. A* **92**, 062315 (2015).

84. Wiebe, N. & Granade, C. Efficient bayesian phase estimation. *Phys. Rev. Lett.* **117**, 010503 (2016).

85. O'Brien, T., Tarasinski, B. & Terhal, B. Quantum phase estimation of multiple eigenvalues for small-scale (noisy) experiments. *New J. Phys.* **21** https://iopscience.iop.org/article/10.1088/1367-2630/aafb8e (2019).

86. Sagastizabal, R. et al. Error Mitigation by symmetry verification on a variational quantum Eigensolver. *Phys. Rev. A* **100**, 010302(R) (2019).

87. Saira, O.-P. et al. Entanglement genesis by ancilla-based parity measurement in 2D circuit QED. *Phys. Rev. Lett.* **112**, 070502 (2014).

88. Fu, X. et al. eQASM: an executable quantum instruction set architecture. In *Proceedings of 25th IEEE International Symposium on High-Performance Computer Architecture (HPCA)*, 224–237 (IEEE, 2019).

89. Johnson, A. & Ungaretti, G. et al. QCoDeS. https://github.com/QCoDeS/Qcodes (2016).
90. Rol, M. A. et al. PycQED. https://github.com/DiCarloLab-Delft/PycQED_py3 (2016).

AUTHOR CONTRIBUTIONS

New quantum algorithms were designed by T.E.O. Errors were estimated by T.E.O., B.S., and A.D. Experiment was performed by R.S. and L.D.C. Density matrix simulations were performed by X.B.M. Computational chemistry calculations performed by B.S. and L.V. Gradient descent and polarizability calculations from experimental and simulated data were performed by T.E.O. and B.S. Relevance to quantum chemistry and suggested applications provided by F.B. and L.V. Paper written by T.E.O., B.S., F.B., and L.V. (with input from all other authors).

COMPETING INTERESTS

The authors declare no competing interests.

ADDITIONAL INFORMATION

Supplementary information is available for this paper at https://doi.org/10.1038/s41534-019-0213-4.

Correspondence and requests for materials should be addressed to T.E.O.

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

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