Bayesian optimisation for variational quantum eigensolvers

B. Moseley* and M. Osborne
Centre for Doctoral Training in Autonomous Intelligent Machines and Systems
University of Oxford

S. Benjamin
Department of Materials
University of Oxford

(Dated: September 14, 2018)

We investigate the performance of Bayesian optimisation for finding the ground state energy of molecules with variational quantum eigensolvers. We implement a variational quantum eigensolver using a classical simulation of a quantum computer and search for the ground state energy of H_2 and LiH molecules. We use the UCC ansatz circuit for H_2 and a hardware-efficient ansatz circuit for LiH. When using a noiseless quantum computer simulation, Bayesian optimisation converges to the ground state energy of H_2 and LiH in an order of magnitude less number of quantum computations than the simultaneous perturbation stochastic approximation (SPSA) algorithm. When simulating a quantum computer with a 10% chance of random gate errors, Bayesian optimisation converges to the ground state ansatz parameter values of H_2 with a significantly higher probability of 83% compared to 18% for SPSA. We suggest Bayesian optimisation is a valuable tool for implementing variational quantum eigensolvers on noisy near-term quantum computers.

I. INTRODUCTION

Quantum algorithms are able to simulate quantum systems, such as molecules, exponentially quicker than classical algorithms [1, 2]. Indeed, it is widely believed that this improvement will yield unprecedented advances in quantum chemistry [3, 4].

Whilst efficient algorithms exist, implementing these algorithms on today's quantum hardware is challenging. Current quantum computers suffer heavily from decoherence and inaccurate gate execution and these sources of noise strongly limit the size of the quantum circuits which can be computed. Two-qubit gate fidelities have been demonstrated in range of 0.1% to 10% and any algorithm which runs on near-term quantum computers must be able cope with these limitations [5].

One strategy for overcoming noise in quantum computers is to design fault-tolerant algorithms, where ancillary qubits are used to monitor and correct the state of a quantum computer [6–8]. Whilst these approaches offer strong theoretical guarantees, they can require a large amount of additional resource, typically in the form of thousands or more of ancillary qubits [9–11]. Near-term quantum computers are likely to have access to tens of qubits, which puts many fault-tolerant algorithms out of reach for near-term quantum computers [12].

Variational quantum eigensolvers (VQEs), first demonstrated in 2014 [13], propose a promising alternative strategy for implementing quantum simulation on current quantum hardware. VQEs are designed to find the ground state energy of a physical system, such as that of a molecule. Importantly, they propose a hybrid classical-quantum approach to overcome noise.

To find the ground state energy, VQEs use optimisation to minimise the expected energy of the system. A quantum computer is used to evaluate the expected energy, which is classically intractable, whilst the high-level optimisation routine is offloaded to a classical computer. By using the quantum computer inside a subroutine, smaller circuits can be used and qubits can be reinitialised after each step, making VQEs more accessible to near-term quantum computers. Recent work by Hempel et al. [14] was able to find the ground state energy of H_2 and LiH using a scalable VQE design on an ion-trap quantum computer.

A key choice in the success of VQEs is the type of classical optimisation algorithm used. In particular, the choice of ansatz used to produce trial states significantly affects the difficulty of the optimisation problem, which is not guaranteed to converge [14–16].

Another key limitation is that VQEs, whilst reducing the requirement on the quantum computer, are still susceptible to noise. Hempel et al. were not able to achieve chemical precision due to quantum hardware noise. Realising VQEs with larger circuit sizes in the presence of noise presents a significant challenge [14].

Bayesian optimisation is a global, classical optimisation method which appears attractive for VQEs. Bayesian optimisation uses the combination of a probabilistic surrogate model and an acquisition function to minimise an objective function. During each optimisation step, objective function evaluations are used to update the surrogate model and acquisition function, and the acquisition function is searched over to find the next objective parameter values for evaluation [17].

A key feature of Bayesian optimisation is that the surrogate model is a probabilistic model, which allows it to accommodate stochastic objective values. Bayesian optimisation is also a global optimisation method. Both of

* bmoseley@robots.ox.ac.uk

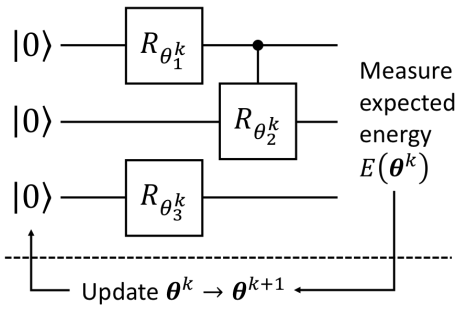

FIG. 1. A VQE uses optimisation to minimise the expected energy of a physical system. This relies on the Ritz-Rayleigh variational principle, which implies that the ground state wavefunction of a system can be found by minimising the expected energy of the system. During optimisation, a quantum computer is used to find the expected energy of a parameterised trial wavefunction. To construct the trial wavefunction, a parameterised ansatz circuit is used. This ansatz circuit may or may not be chemistry-inspired. The expected energy of the system is estimated by making repeated measurements of the trial state, and handed to a classical optimisation routine which updates the parameters of the trial wavefunction. Dashed line separates quantum computation (top) from classical computation (bottom).

these considerations are important when implementing VQEs on noisy near-term quantum computers where the true ground state energy is desired.

Existing implementations of VQEs have mostly employed the Nelder-Mead algorithm or approximate gradient methods for classical optimisation [13, 14, 18]. In this work we quantitatively assess the performance of Bayesian optimisation.

We assess the performance of Bayesian optimisation for finding the ground state energy of H_2 and LiH with VQEs. We also investigate the convergence of Bayesian optimisation for finding the ground state energy of H_2 with increasing levels of gate noise in the quantum computer. We use a classical simulation of a quantum computer to carry out our tests. We compare all our results to the simultaneous perturbation stochastic approximation (SPSA) optimisation algorithm, which has been used in previous VQE demonstrations [18].

II. METHODS

The VQE method is summarised in Figure 1. To assess the performance of Bayesian optimisation, we build a VQE using a classical simulation of a quantum computer.

We design our simulator using the open-source ProjectQ library [19]. Our simulator accepts an ansatz circuit, a set of ansatz parameters values and a Hamiltonian representing a physical system as input. The output of our simulator is an estimate of the expected energy of the system.

To map molecular Hamiltonians onto a quantum computer, they are typically re-written in the second quantised form and transformed using a suitable qubit mapping, such as the Jordan-Wigner transformation or Bravyi-Kitaev transformation [14, 20]. Both transformations produce qubit Hamiltonians which are a linear combinations of tensor products of Pauli operators and our simulator accepts any Hamiltonian which can be expressed in this form.

To estimate the expected energy, we initialise our simulator in the $|0\rangle^{\otimes n}$ state, where n is the number of qubits in the ansatz circuit. We prepare a trial wavefunction by executing the ansatz circuit, using the input ansatz parameter values. Given that the Hamiltonian is a linear combination of Pauli terms, we estimate the expected energy by averaging repeated measurements of the trial state in the Pauli basis for each term in the Hamiltonian and summing the results. This estimation method produces shot noise on the expected energy for a finite number of measurements [21]. When using our simulator it is also possible to classically obtain the exact expected value of the Hamiltonian as we have access to the simulator's hidden state amplitudes.

To make the simulation more realistic, we simulate gate noise during the execution of the ansatz circuit. For each gate we allow the chance that a different gate is randomly executed, and we define the gate noise as the probability a different gate is executed to the gate specified. For single qubit gates, if a different gate is executed it is chosen randomly from the set $\{X_1, Y_1, Z_1\}$. For two qubit gates, it is chosen randomly from the set $\{X_1, X_2, Y_1, Y_2, Z_1, Z_2, X_1X_2, X_1Y_2, Y_1X_2, X_1Z_2, Z_1X_2, Y_1Y_2, Y_1Z_2, Z_1Y_2, Z_1Z_2\}$.

To carry out Bayesian optimisation, we use the GPyOpt library [22]. For all our tests we use a Gaussian process for our surrogate model with a Matern-$\frac{5}{2}$ kernel. We update our surrogate model and propose a new set of ansatz parameter values after each expected energy estimate. We optimise for a fixed number of steps before testing for convergence.

We compare all our results to the SPSA optimisation algorithm, an approximate gradient descent method which can accommodate noise in the objective function [23]. We fix the SPSA look size, step size, alpha and gamma values to be 0.05 radians, 0.05 radians, 0.3 and 0.2 respectively.

For our first test, we search for the ground state energy of H_2 and LiH without the presence of noise. To simulate H_2, we use the 4 qubit Jordan-Wigner transformed Hamiltonian described by Hempel et al., at an inter-nuclear distance of $R = 0.75\mathring{A}$. The corresponding ground state energy at this distance is -1.1371 Hartree. We use the chemistry-motivated Unitary Coupled Cluster (UCC) ansatz [24], shown in Fig 5, which has 3 free parameters.

To simulate LiH, we use the 6 qubit Bravyi-Kitaev

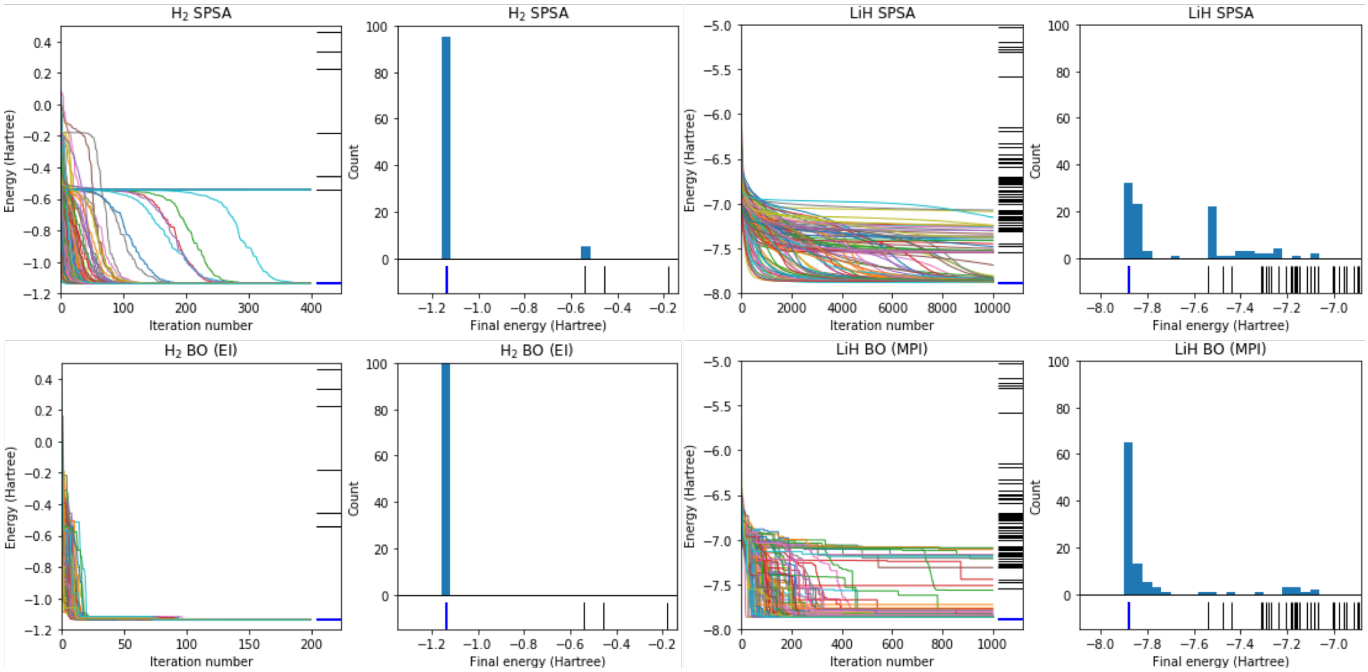

FIG. 2. Convergence of Bayesian optimisation for finding the ground state energy of H_2 and LiH without the presence of noise. Left: convergence for H_2. Right: convergence for LiH. Top row shows a comparison to the SPSA optimisation algorithm. Inner left shows the current lowest expected energy value against iteration number for 100 different randomly-initialised optimisation runs. Inner right shows the histogram of final expected energy values over all 100 runs. The energy eigenspectrum of each Hamiltonian is shown in each plot. Black lines show excited states, thick blue line shows the true ground state energy. We show results for the EI acquisition function for H_2 and the MPI acquisition function for LiH.

transformed Hamiltonian described by McArdle et al. [25] at an inter-nuclear distance of $R = 1.45\mathring{A}$. The corresponding ground state energy at this distance is -7.8807 Hartree. We also use their hardware-efficient ansatz, shown Fig 5, which has 42 free parameters. With a larger number of parameters, this presents a more challenging optimisation problem. For both tests we do not include gate noise and use the exact expected energy values without shot noise from our simulator. We test the convergence of the expected improvement (EI), maximum probability of improvement (MPI) and lower confidence bound (LCB) acquisition functions for Bayesian optimisation [22].

For our second test, we search for the ground state energy of H_2 in the presence of shot noise and varying levels of gate noise, using the estimated expected energy values from our simulator. We use the same ansatz and Hamiltonian as our first test. We keep the shot noise fixed, averaging 100 measurements of each term in the Hamiltonian to estimate the expected energy. We test the convergence of both SPSA and Bayesian optimisation for different levels of gate noise ranging from 0% to 10%, in increments of 1%. In this test we use the EI acquisition function for Bayesian optimisation.

Our second test requires considerably more computational power, as we average over 100 preparations and measurements of the trial state to estimate the expected energy of each Hamiltonian term. We run expected energy measurements in parallel using 15 CPU cores, resulting in each expected energy estimate taking approximately 4 seconds to compute.

Finally, for all tests we re-run each optimisation method 100 times with a different random initialisation each time to gather statistics on the likelihood of convergence.

III. RESULTS

Figure 2 shows the convergence of Bayesian optimisation without the presence of noise. For H_2, we find all 100 Bayesian optimisation runs using the EI acquisition function converge to within 0.06 Hartree of the ground state energy, after approximately 20 evaluations of the expected energy. We observe similar convergence behaviour across all three types of acquisition function tested. For SPSA, we find 95 of the runs converge to within 0.06 Hartree of the ground state energy after approximately 200 evaluations (we note 2 evaluations of the expected energy are required per iteration to compute the SPSA gradient). The remaining 5 SPSA runs converge to the first excited state eigenvalue.

For LiH, we find 78 of the 100 Bayesian optimisation runs using the MPI acquisition function converge

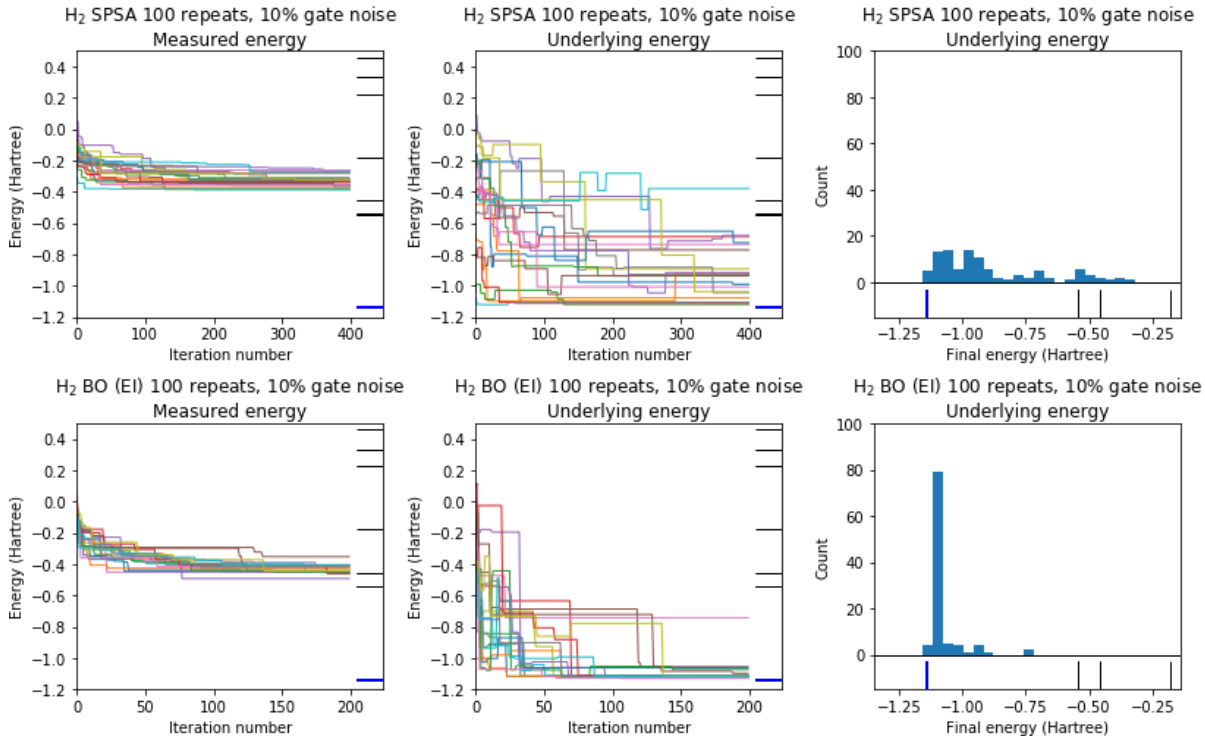

FIG. 3. Convergence of Bayesian optimisation for finding the ground state energy of H_2 in the presence of shot noise and 10% gate noise. Top row shows a comparison to the SPSA optimisation algorithm. Left shows the current lowest estimated expected energy value against iteration number for 100 different randomly-initialised optimisation runs. Middle shows the underlying exact expected energy value corresponding to the current best ansatz parameter values against iteration number for each of the 100 runs, without noise, calculated classically using the hidden state of the simulator. Right shows the histogram of final values of these underlying exact expected energy values over all 100 runs. The energy eigenspectrum is shown in each plot. Black lines show excited states, thick blue line shows the true ground state energy. The EI acquisition function is used for Bayesian optimisation.

to within 0.06 Hartree of the ground state energy, after approximately 500 evaluations. When using the LCB and EI acquisition functions, we find 76 and 43 of the runs converge respectively, in a similar number of evaluations. For SPSA, we find 55 of runs converge to within 0.06 Hartree of the ground state energy, after 20,000 expected energy evaluations.

Figure 3 shows the convergence of Bayesian optimisation in the presence of shot noise and 10% gate noise. We plot the convergence of both the estimated expected energy seen by the VQE and the underlying exact expected energy corresponding to the current best ansatz parameter values, calculated using the hidden state amplitudes in our simulator. We find 83 of the Bayesian optimisation runs converge to ansatz parameter values which correspond to an underlying exact expected energy value within 0.06 Hartree of the ground state energy, after approximately 150 evaluations. For SPSA, 18 of the runs converge under the same convergence condition after 800 evaluations.

Figure 4 shows the probability these runs will converge for varying levels of gate noise. We use the same convergence condition described above to calculate this probability. We find Bayesian optimisation converges more than 80% of the time for all levels of gate noise up to 10%, whilst the probability of convergence for the SPSA algorithm degrades strongly below 80% after approximately 4% gate noise.

IV. CONCLUSIONS AND LIMITATIONS

We have assessed the performance of Bayesian optimisation for finding the ground state energy of H_2 and LiH molecules with variational quantum eigensolvers. We used a classical simulation of a quantum computer to implement our variational quantum eigensolver.

Our work suggests that Bayesian optimisation could be a valuable tool for implementing VQEs on noisy near-term quantum computers. Without the presence of noise, we found Bayesian optimisation converged to the ground state energy of H_2 and LiH in an order of magnitude less number of quantum computations than the SPSA algorithm. In the presence of shot noise and gate noise, Bayesian optimisation converged to the ground state energy of H_2 more reliably than the SPSA algorithm.

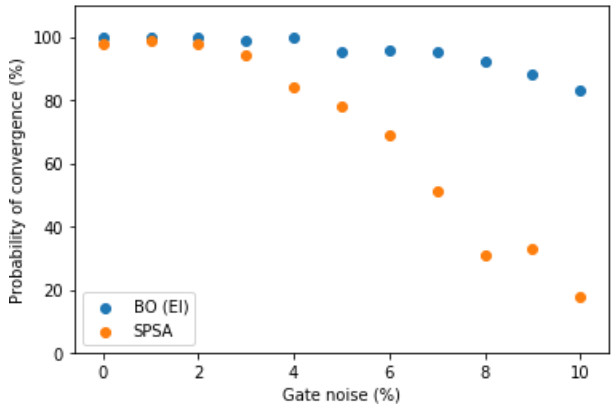

FIG. 4. Probability a run will converge to ansatz parameter values which correspond to an underlying exact expected energy value within 0.06 Hartree of the ground state energy of H_2, for varying levels of gate noise. Blue shows the probability for Bayesian optimisation, orange shows the probability for the SPSA algorithm.

Bayesian optimisation was able to find ansatz parameter values which corresponded to exact expected energy values close to the true ground state of H_2, even though the noisy estimated expected energy values seen during optimisation were significantly above the ground state energy (Figure 3, left). We suspect the higher values were due to noise in the VQE pushing the average of each Pauli measurement in the Hamiltonian closer to 0 when estimating the expected energy.

Much further work could be carried out to extend these results. The performance of Bayesian optimisation could be assessed for a wider variety of ansatze and Hamiltonians, including those of more complex molecules. Other acquisition functions could be tested, particularly ones which explicitly handle noise in the objective function [17, 26]. We note that the EI acquisition function performed poorly compared to the MPI and LCB acquisition functions when searching for the ground state of LiH using the hardware-efficient ansatz. Bayesian optimisation could also be tested with varying levels of shot noise, which we were not able to carry out due to computational constraints.

Whilst showing good convergence to the local vicinity of the ground state parameter values, Bayesian optimisation did not often achieve high precision on the ground state energy value. Using a hybrid approach by switching to gradient descent close to the solution may improve this [27, 28]. The performance of Bayesian optimisation could also be compared to other optimisation methods, such as the Nelder-Mead algorithm, and other variational techniques, such as the imaginary time method [25].

Finally, our preliminary results could be confirmed experimentally on current quantum hardware.

ACKNOWLEDGMENTS

We would like to thank Sam McArdle for his extensive advice on VQE theory and his VQE code from which this work was built upon, as well as the rest of the Quantum Technology Theory Group at Oxford University for their advice.

We would like to thank the ProjectQ group for providing an open-source framework for quantum computing and the GPyOpt group for their Bayesian optimisation library.

[1] Alán Aspuru-Guzik, Anthony D. Dutoi, Peter J. Love, and Martin Head-Gordon, "Simulated quantum computation of molecular energies," Science **309**, 1704–1707 (2005).

[2] Ivan Kassal, James D. Whitfield, Alejandro Perdomo-Ortiz, Man-Hong Yung, and Alán Aspuru-Guzik, "Simulating chemistry using quantum computers," Annual Review of Physical Chemistry **62**, 185–207 (2011).

[3] Alán Aspuru-Guzik, Roland Lindh, and Markus Reiher, "The matter simulation (r)evolution," ACS Central Science **4**, 144–152 (2018).

[4] B. P. Lanyon, J. D. Whitfield, G. G. Gillett, M. E. Goggin, M. P. Almeida, I. Kassal, J. D. Biamonte, M. Mohseni, B. J. Powell, M. Barbieri, A. Aspuru-Guzik, and A. G. White, "Towards quantum chemistry on a quantum computer," Nature Chemistry **2**, 106 EP – (2010).

[5] Amir Fruchtman and Iris Choi, "Technical roadmap for fault-tolerant quantum computing," NQIT Technical Roadmap (2016).

[6] Austin G. Fowler, Matteo Mariantoni, John M. Martinis, and Andrew N. Cleland, "Surface codes: Towards practical large-scale quantum computation," Phys. Rev. A **86** (2012).

[7] D. Gottesman, "Fault-Tolerant Quantum Computation," ArXiv e-prints (2007).

[8] P. W. Shor, "Fault-tolerant quantum computation," in *Proceedings of 37th Conference on Foundations of Computer Science* (1996) pp. 56–65.

[9] Markus Reiher, Nathan Wiebe, Krysta M. Svore, Dave Wecker, and Matthias Troyer, "Elucidating reaction mechanisms on quantum computers," Proceedings of the National Academy of Sciences **114**, 7555–7560 (2017).

[10] Masoud Mohseni, Peter Read, Hartmut Neven, Sergio Boixo, Vasil Denchev, Ryan Babbush, Austin Fowler, Vadim Smelyanskiy, and John Martinis, "Commercialize quantum technologies in five years," Nature **543**, 171 (2017).

[11] A. Paler and S. J. Devitt, "An introduction into fault-tolerant quantum computing," in *2015 52nd ACM/EDAC/IEEE Design Automation Conference (DAC)* (2015).

[12] J. Preskill, "Quantum Computing in the NISQ era and beyond," ArXiv e-prints (2018).

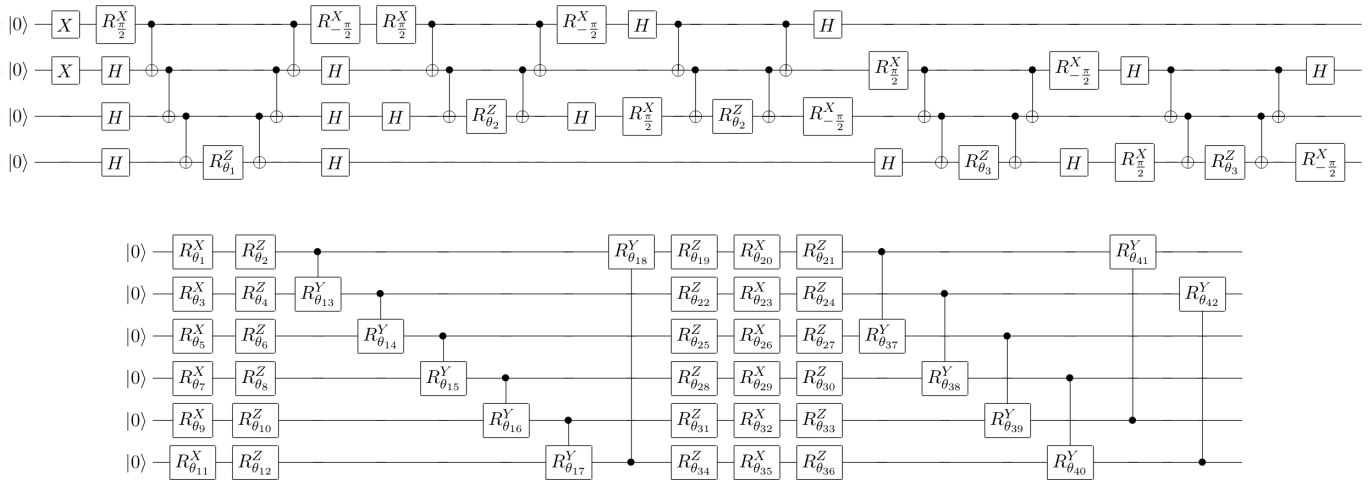

FIG. 5. Ansatz circuits used in our VQE simulations. Top: H_2 UCC ansatz circuit with 4 qubits and 3 free parameters. Bottom: LiH hardware-efficient ansatz circuit with 6 qubits and 42 free parameters.

[13] 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).

[14] C. Hempel, C. Maier, J. Romero, J. McClean, T. Monz, H. Shen, P. Jurcevic, B. P. Lanyon, P. Love, R. Babbush, A. Aspuru-Guzik, R. Blatt, and C. F. Roos, "Quantum Chemistry Calculations on a Trapped-Ion Quantum Simulator," Physical Review X **8** (2018).

[15] J. R. McClean, S. Boixo, V. N. Smelyanskiy, R. Babbush, and H. Neven, "Barren plateaus in quantum neural network training landscapes," ArXiv e-prints (2018).

[16] Dave Wecker, Matthew B. Hastings, and Matthias Troyer, "Progress towards practical quantum variational algorithms," Phys. Rev. A **92** (2015).

[17] P. I. Frazier, "A Tutorial on Bayesian Optimization," ArXiv e-prints (2018).

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

[19] D. S. Steiger, T. Häner, and M. Troyer, "ProjectQ: An Open Source Software Framework for Quantum Computing," ArXiv e-prints (2016).

[20] Jacob T. Seeley, Martin J. Richard, and Peter J. Love, "The bravyi-kitaev transformation for quantum computation of electronic structure," The Journal of Chemical Physics **137**, 224109 (2012).

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

[22] The GPyOpt authors, "Gpyopt: A bayesian optimization framework in python," `http://github.com/SheffieldML/GPyOpt` (2016).

[23] J. C. Spall, "Multivariate stochastic approximation using a simultaneous perturbation gradient approximation," IEEE Transactions on Automatic Control **37**, 332–341 (1992).

[24] J. Romero, R. Babbush, J. R. McClean, C. Hempel, P. Love, and A. Aspuru-Guzik, "Strategies for quantum computing molecular energies using the unitary coupled cluster ansatz," ArXiv e-prints (2017).

[25] S. McArdle, S. Endo, T. Jones, Y. Li, S. Benjamin, and X. Yuan, "Variational quantum simulation of imaginary time evolution with applications in chemistry and beyond," ArXiv e-prints (2018).

[26] J. M. Hernández-Lobato, M. W. Hoffman, and Z. Ghahramani, "Predictive Entropy Search for Efficient Global Optimization of Black-box Functions," ArXiv e-prints (2014).

[27] Binxin Ru, Mark McLeod, Diego Granziol, and Michael A. Osborne, "Fast information-theoretic bayesian optimisation," in *International Conference on Machine Learning (ICML)* (2018).

[28] J. Wu, M. Poloczek, A. G. Wilson, and P. I. Frazier, "Bayesian Optimization with Gradients," ArXiv e-prints (2017).