Variational quantum simulation of imaginary time evolution with applications in chemistry and beyond

Sam McArdle,[1] Suguru Endo,[1] Ying Li,[2] Simon Benjamin,[1,*] and Xiao Yuan[1,†]

[1] *Department of Materials, University of Oxford, Parks Road, Oxford OX1 3PH, United Kingdom*
[2] *Graduate School of China Academy of Engineering Physics, Beijing 100193, China*

(Dated: May 21, 2019)

Imaginary time evolution is a powerful tool in the study of many-body quantum systems. While it is conceptually simple to simulate such evolution with a classical computer, the time and memory requirements scale exponentially with the system size. Conversely, quantum computers can efficiently simulate many-body quantum systems, but the non-unitary nature of imaginary time evolution is incompatible with canonical unitary quantum circuits. Here we propose a variational method for simulating imaginary time evolution on a quantum computer, using a hybrid algorithm that combines quantum and classical resources. We apply this technique to the problem of finding the ground state energy of many-particle Hamiltonians. We numerically test our algorithm on problems in quantum computational chemistry; specifically, finding the ground state energy of the Hydrogen molecule and Lithium Hydride. Our algorithm successfully finds the global ground state with high probability, outperforming gradient descent optimisation, which commonly becomes trapped in local minima. Our method can also be applied to general optimisation problems, Gibbs state preparation, and quantum machine learning. As our algorithm is hybrid, suitable for error mitigation methods, and can exploit shallow quantum circuits, it can be implemented with current and near-term quantum computers.

arXiv:1804.03023v1 [quant-ph] 9 Apr 2018

Imaginary time is an unphysical, yet powerful, mathematical concept. It has been utilised in numerous physical domains including: quantum mechanics, statistical mechanics, and cosmology. Often referred to as performing a 'Wick rotation' [1], replacing real time with imaginary time connects Euclidean and Minkowski space [2], quantum and statistical mechanics [3], and static problems to problems of dynamics [4]. In quantum mechanics, propagating a wavefunction in imaginary time enables: the study of the finite temperature properties of many-body quantum systems [5–7], finding the ground state wavefunction and energy [8–11], and simulating real time dynamics [12, 13]. For a system with Hamiltonian, H, evolving in real time, t, the propagator is given by e^{-iHt}. The corresponding propagator in imaginary time, $\tau = it$, is given by $e^{-H\tau}$. We note that while the real time propagator is a unitary operator, the imaginary time propagator is not.

Using a classical computer, we can simulate this (non-unitary) imaginary time evolution by evaluating the propagator and applying it to the system wavefunction. However, because the dimension of the wavefunction grows exponentially with the number of particles, classical simulation of many-body quantum systems is limited to small or specific cases [14]. While efficient variational trial states have been developed for a number of applications [15], powerful trial wavefunctions typically require classical computational resources which scale exponentially with the system size [11].

Quantum computing can naturally and efficiently store many-body quantum states, and hence is suitable for simulating quantum systems [16]. We can map the system Hamiltonian to a qubit Hamiltonian, and simulate real time evolution (as described by the Schrödinger equation) by realising the corresponding unitary evolution with a quantum circuit [17]. Using Trotterization [18], the real time propagator can be decomposed into a sequence of single and two qubit gates [19]. The ability to represent the real time propagator with a sequence of gates stems from its unitarity. In contrast, because the imaginary time operator is non-unitary, it is impossible to decompose it into a sequence of unitary gates using Trotterization, and thus directly realise it with a quantum circuit. As a result, alternative methods must be used to carry out imaginary time evolution using a quantum computer.

Classically, we can simulate real (imaginary) time evolution of parametrised trial states by repeatedly solving the (Wick-rotated) Schrödinger equation over a small timestep, and updating the parameters for the next timestep [8, 9, 11, 20–23]. This method has recently been extended to quantum computing, where it was used to simulate real time dynamics [24]. Closely related are the variational quantum eigensolver (VQE) [25–31] and the quantum approximate optimisation algorithm (QAOA) [32], which update the parameters using a classical optimisation routine, to find the minimum energy eigenvalue of a given Hamiltonian. As 'hybrid quantum-classical methods', these algorithms use a small quantum computer to carry out a classically intractable subroutine, and a classical computer to solve the higher level problem. The quantum subroutine may only require a small number of qubits and a low depth circuit, presenting a potential use for noisy intermediate-scale quantum hardware [33].

In this letter, we propose a method to simulate imag-

inary time evolution on a quantum computer, using a hybrid quantum-classical variational algorithm. The proposed method thus combines the power of quantum computers to efficiently represent many-body quantum states, with classical computers' ability to simulate arbitrary (including unphysical) processes. We discuss using this method to find the ground state energy of many-body quantum systems, and to solve optimisation problems. We then numerically test the performance of our algorithm at finding the ground state energy of both the Hydrogen molecule and Lithium Hydride. As this simulation method only requires a low depth circuit, it can be realised with current and near-term quantum processors.

Variational simulation of imaginary time evolution.— We focus on many-body systems that are described by Hamiltonians $H = \sum_i \lambda_i h_i$, with real coefficients, λ_i, and observables, h_i, that are tensor products of Pauli matrices. We assume that the number of terms in this Hamiltonian scales polynomially with the system size, which is true for many physical systems, such as molecules [34]. Given an initial state $|\psi\rangle$, the normalised imaginary time evolution is defined by

$$|\psi(\tau)\rangle = A(\tau)e^{-H\tau}\,|\psi(0)\rangle, \tag{1}$$

where $A(\tau) = 1/\sqrt{\langle\psi(0)|\,e^{-2H\tau}\,|\psi(0)\rangle}$ is a normalisation factor. In the instance that the initial state is a maximally mixed state, the state at time τ is a thermal or Gibbs state $\rho_{T=1/\tau} = e^{-H\tau}/\mathrm{Tr}[e^{-H\tau}]$, with temperature $T = 1/\tau$. When the initial state has a non-zero overlap with the ground state, the state at $\tau \to \infty$ is the ground state of H. Evolution under Eq. (1) corresponds to the Wick rotated Schrödinger equation,

$$\frac{\partial\,|\psi(\tau)\rangle}{\partial\tau} = -(H - E_\tau)\,|\psi(\tau)\rangle\,, \tag{2}$$

where the term $E_\tau = \langle\psi(\tau)|H|\psi(\tau)\rangle$ results from enforcing normalisation. Even if $|\psi(\tau)\rangle$ could be represented by a quantum computer, the non-unitary imaginary time evolution cannot be naively mapped to a quantum circuit.

In our variational method, instead of directly encoding the quantum state $|\psi(\tau)\rangle$ at time τ, we approximate it using a parametrised trial state $|\phi(\vec{\theta}(\tau))\rangle$, with $\vec{\theta}(\tau) = (\theta_1(\tau), \theta_2(\tau), \ldots, \theta_N(\tau))$. This stems from the intuition that the physically relevant states are contained in a small subspace of the full Hilbert space [35]. The trial state is referred to as the ansatz. In condensed matter physics and computational chemistry, a wide variety of ansatze have been proposed for both classical and quantum variational methods [11, 16, 36, 37].

Using a quantum circuit, we prepare the trial state, $|\phi(\vec{\theta})\rangle$, by applying a sequence of parametrised unitary gates, $V(\vec{\theta}) = U_N(\theta_N)\ldots U_k(\theta_k)\ldots U_1(\theta_1)$ to our initial state, $|\bar{0}\rangle$. We express this as $|\phi(\vec{\theta})\rangle = V(\vec{\theta})\,|\bar{0}\rangle$ and remark that $V(\vec{\theta})$ is also referred to as the ansatz. We refer to all possible states that could be created by the circuit V as the 'ansatz space'. Here, $U_k(\theta_k)$ is the k^{th} unitary gate, controlled by parameter θ_k, and the gate can be regarded as a single or two qubit gate.

To simulate the imaginary time evolution of the trial state, we use McLachlan's variational principle [38],

$$\delta\|(\partial/\partial\tau + H - E_\tau)\,|\psi(\tau)\rangle\,\| = 0, \tag{3}$$

where $\|\rho\| = \mathrm{Tr}[\sqrt{\rho\rho^\dagger}]$ denotes the trace norm of a state. By replacing $|\psi(\tau)\rangle$ with $|\phi(\tau)\rangle = |\phi(\vec{\theta}(\tau))\rangle$, we effectively project the desired imaginary time evolution onto the manifold of the ansatz space. The evolution of the parameters is obtained from the resulting differential equation

$$\sum_j A_{ij}\dot{\theta}_j = C_i, \tag{4}$$

where

$$\begin{aligned} A_{ij} &= \Re\left(\frac{\partial\,\langle\phi(\tau)|}{\partial\theta_i}\frac{\partial\,|\phi(\tau)\rangle}{\partial\theta_j}\right), \\ C_i &= \Re\left(-\sum_\alpha \lambda_\alpha \frac{\partial\,\langle\phi(\tau)|}{\partial\theta_i} h_\alpha\,|\phi(\tau)\rangle\right). \end{aligned} \tag{5}$$

As both A_{ij} and C_i are real, the derivative $\dot{\theta}_j$ is also real, as required for parametrising a quantum circuit. Interestingly, although the average energy term E_τ appears in Eq. (2), it does not appear in Eq. (4). This is because the ansatz applied maintains normalisation, as it is composed of unitary operators. The derivation of Eq. (4) can be found in the Supplementary Materials.

By following a similar method to that introduced in Refs. [24, 39, 40], we can efficiently measure A_{ij} and C_i using a quantum computer. Assume that the derivative of a unitary gate $U_i(\theta_i)$ can be expressed as $\partial U_i(\theta_i)/\partial\theta_i = \sum_k f_{k,i} U_i(\theta_i)\sigma_{k,i}$, with unitary operator $\sigma_{k,i}$. The derivative of the trial state is given by $\partial\,|\phi(\tau)\rangle/\partial\theta_i = \sum_k f_{k,i}\tilde{V}_{k,i}\,|\bar{0}\rangle$, with $\tilde{V}_{k,i} = U_N(\theta_N)\ldots U_{i+1}(\theta_{i+1})U_i(\theta_i)\sigma_{k,i}\ldots U_1(\theta_1)$. There are typically only one or two terms resulting from each derivative. As an example, when $U_i(\theta_i)$ is a single qubit rotation $R_z(\theta_i) = e^{-i\theta_i\sigma_z/2}$, the derivative $\partial U_i(\theta_i)/\partial\theta_i = -i/2\times\sigma_z e^{-i\theta_i\sigma_z/2}$. The coefficients $A_{i,j}$ and C_i are given by

$$\begin{aligned} A_{i,j} &= \Re\left(\sum_{k,l} f^*_{k,i} f_{l,j}\,\langle\bar{0}|\,\tilde{V}^\dagger_{k,i}\tilde{V}_{l,j}\,|\bar{0}\rangle\right), \\ C_i &= \Re\left(\sum_{k,\alpha} f^*_{k,i}\lambda_\alpha\,\langle\bar{0}|\,\tilde{V}^\dagger_{k,i} h_\alpha V\,|\bar{0}\rangle\right). \end{aligned} \tag{6}$$

All of these terms are of the form $a\Re(e^{i\theta}\,\langle\bar{0}|\,U\,|\bar{0}\rangle)$ and can be evaluated using the circuit in Fig. 1. In practice, we do

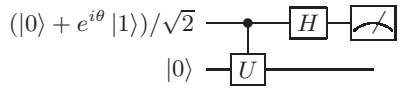

FIG. 1. A quantum circuit that evaluates $\Re(e^{i\theta}\langle\bar{0}|U|\bar{0}\rangle)$. H is the Hadamard gate.. The first qubit is measured in the computational $\{|0\rangle, |1\rangle\}$ basis, and the average value $\langle Z\rangle$ of the Pauli Z matrix equals the $\Re(e^{i\theta}\langle\bar{0}|U|\bar{0}\rangle)$.

not need to realise the controlled-U gate, and can instead use a more simple circuit. We refer to the Supplementary Materials and Ref. [24] for details.

With $A(\tau)$ and $C(\tau)$ at time τ, the imaginary time evolution over a small interval $\delta\tau$ can be simulated by evaluating $\dot{\vec{\theta}}(\tau)$, and using a suitable update rule, such as the Euler method,

$$\begin{aligned}\vec{\theta}(\tau+\delta\tau) &= \vec{\theta}(\tau) + \delta\tau\cdot\dot{\vec{\theta}}(\tau),\\ &= \vec{\theta}(\tau) + A^{-1}(\tau)\cdot C(\tau)\delta\tau.\end{aligned} \tag{7}$$

By repeating this process $N_T = \tau_{total}/\delta\tau$ times, we can simulate imaginary time evolution over a duration τ_{total}. When there are redundant parameters, the rank of A can be less than the number of parameters, which suggests a non-unique solution for Eq. (4). This can be circumvented by fixing the values of some of the parameters until A becomes invertible [11].

Ground state energy via imaginary time evolution.— We apply our method to the problem of finding the ground state energy of a many-body Hamiltonian, H. As with the VQE, our goal is to find the values of the parameters, $\vec{\theta}$, which minimise the expectation value of the Hamiltonian

$$E_{\min} = \min_{\vec{\theta}} \langle\phi(\vec{\theta})|H|\phi(\vec{\theta})\rangle, \tag{8}$$

where $|\phi(\vec{\theta})\rangle = U_N(\theta_N)\ldots U_2(\theta_2)U_1(\theta_1)|\bar{0}\rangle$ is our variational trial state. The VQE solves this problem by using a quantum computer to construct a good ansatz and measure the expectiation value of the Hamiltonian, and a classical optimisation routine to obtain new values of the parameters. To date, the VQE has been used to experimentally find the ground state energies of several small molecules [25–31]. In order to preserve the exponential speedup of the VQE over classical methods, the trial state is constructed using a number of parameters that scales polynomially with the system size. However, because we may need to consider many possible values for each parameter, the total size of the parameter space still scales exponentially with the system size. Moreover, many optimisation algorithms, such as gradient descent, are liable to becoming trapped in local minima. This combination can make the classical optimisation step of the VQE very difficult [41].

As described above, if the initial state has a non-zero overlap with the ground state, the propagation in imaginary time will evolve the system into the ground state, in the limit that $\tau\to\infty$. Classically, this has been leveraged as a powerful tool to find the ground state energy of quantum systems [8, 9, 11]. Using our method, we can efficiently simulate imaginary time evolution to find the ground state, using a quantum computer. In the numerical simulations described below, we use the Euler method to solve differential equations, which corresponds to the update rule for the parameters shown in Eq. (7). We prove in the Supplementary Materials that when $\delta\tau$ is sufficiently small, the average energy of the trial state, $E(\tau) = \langle\phi(\tau)|H|\phi(\tau)\rangle$, always decreases when following the Euler update rule; $E(\tau+\delta\tau)\leq E(\tau)$.

The update rule for the conventional gradient descent method is given by

$$\vec{\theta}(\tau+\delta\tau) = \vec{\theta}(\tau) + a\vec{G}(\tau)\delta\tau = \vec{\theta}(\tau) + aC(\tau)\delta\tau, \tag{9}$$

where a is a constant scaling factor, $\vec{G}(\tau) = -\nabla E(\tau)$ is the gradient of $E(\tau)$, and $C(\tau)\equiv -\nabla E(\tau)$ is the same vector given in Eq. (5). The distinction between these two update rules is the prescence of the A matrix in Eq. (7). Thus, the gradient descent method only considers information about the average energy, and not about the ansatz itself, which is encoded in A.

Simulation of H_2 and LiH.— We now use our method to find the ground state energy of the H_2 and LiH molecules in their minimal spin orbital basis sets. We first map the fermionic Hamiltonian into a qubit Hamiltonian using the Bravyi-Kitaev encoding [42] and utilise a procedure for Hamiltonian reduction. For details of the chemistry simulation methods we refer the reader to Ref. [37, 43–45]. After the encoding procedure, the two qubit Hamiltonian of H_2 [27] is

$$H = g_0 + g_1Z_0 + g_2Z_1 + g_3Z_0Z_1 + g_4Y_0Y_1 + g_5X_0X_1, \tag{10}$$

where g_i are real coefficients determined by the distance R between the two nuclei, and X, Y, Z denote the Pauli matrices. Here, we consider an internuclear distance of $R = 0.75\mathring{A}$ and hence $g_0 = 0.2252$, $g_1 = 0.3435$, $g_2 = -0.4347$ $g_3 = 0.5716$, $g_4 = 0.0910$, $g_5 = 0.0910$. There are numerous possible choices for the ansatz circuit, and we consider two of these: the unitary coupled cluster (UCC) [37] and hardware efficient [31] ansatze. The UCC ansatz is used to construct a chemically motivated trial state, and is given by

$$|\psi(\theta)\rangle = e^{-i\theta X_0Y_1}|0\rangle_1|1\rangle_0. \tag{11}$$

This trial state can be prepared using the quantum circuit in Fig. 2(a). In contrast, the hardware efficient ansatz does not construct a chemically motivated trial state, but instead attempts to generate a flexible wavefunction using as few gates as possible. There are different ways of designing a hardware efficient ansatz, and the ansatz we use is shown in Fig. 2(c).

As the UCC ansatz has only one parameter θ, both A and C defined in Eq. (5) are scalars. The value of

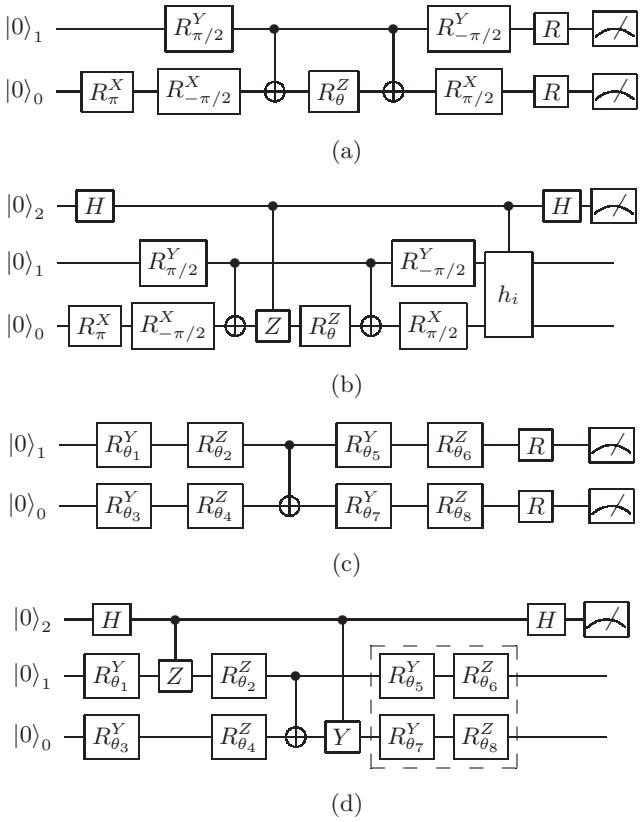

FIG. 2. The quantum circuits for the H_2 molecule in the minimal spin orbital basis set. Measurements are in the Z basis, and rotation gates are applied before measurement to measure in the X or Y basis. (a) The quantum circuit [27] for preparing the UCC ansatz. Here, $R_\alpha^{\sigma_j} = e^{-i\alpha\sigma_j/2}$. The parameter θ is updated in order to minimise the total energy. (b) The quantum circuit that estimates C, as defined in Eq. (5). Here, h_i corresponds to one of the six terms in Eq. (10). As h_i is a tensor product of Pauli matrices, the control-h_i gate can be realised by applying the individual control-Pauli gates in turn. (c) The quantum circuit for preparing the hardware efficient ansatz with eight parameters. In this ansatz, one parameter is redundant. (d) The circuit to measure $A_{2,7} = \Re\left(\frac{\partial\langle\phi(\tau)|}{\partial\theta_2}\frac{\partial|\phi(\tau)\rangle}{\partial\theta_7}\right)$. In practice, the gates in the dashed box may be omitted. The other terms of A and C can be measured using similar circuits.

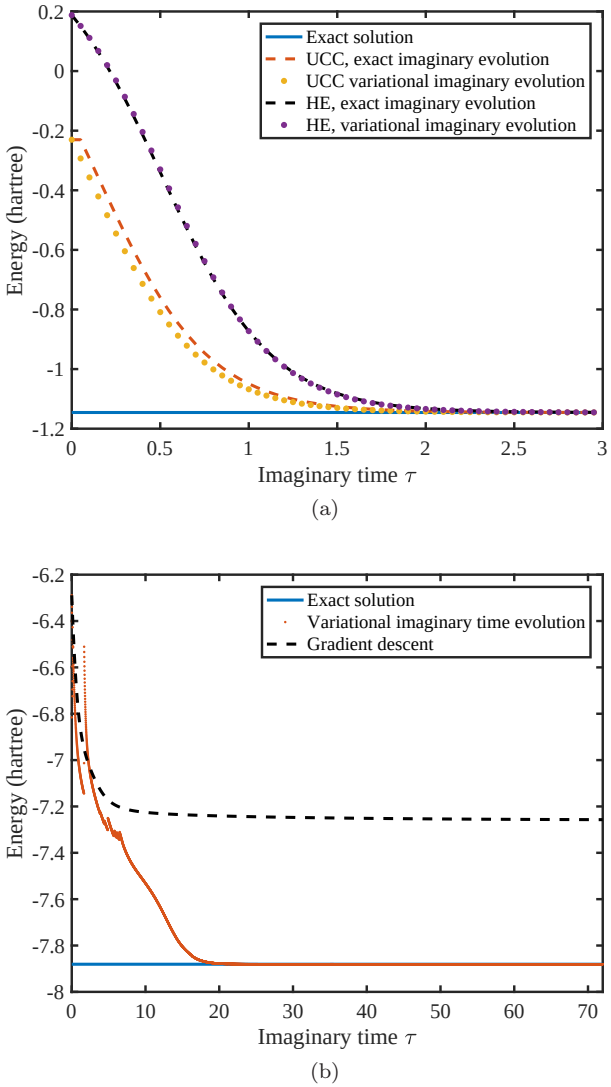

FIG. 3. (Color online) Simulation results for H_2 and LiH with randomly initialised parameters. The blue line is the exact value of the ground state energy. (a) Simulation result for H_2. Time step is $\delta\tau = 0.05$. The dashed lines are the exact simulation of the imaginary time evolution in Eq. (2). The dot points are simulation results of the variational imaginary time evolution method. UCC: unitary coupled cluster; HE: hardware efficient. (b) Simulation result for LiH. Time step is $\delta\tau = 0.01$. The red dot points are the simulation results of the variational imaginary time evolution method, which finds the correct ground state energy. The black dashed line is the result using the gradient descent method, which converges to a local minimum. The discontinuities in the imaginary time evolution result from the large value of $\delta\tau$ used.

C can be measured using the circuit in Fig. 2(b), and $A = 0.25$. As A is a constant, the imaginary time evolution is equivalent to the gradient descent method, with the gradient calculated by a quantum circuit [39, 40]. The optimal value of the parameter is $\theta_{\text{opt}} = 5.2146$ and the minimum energy is $E_{\text{min}} = -1.1457$ Hartree. When considering the hardware efficient ansatz, we have eight parameters, and A is no longer a scalar. The simulation results of the variational imaginary time evolution are shown in Fig. 3(a). We see that the variational imaginary time evolution is able to converge to the ground state energy for both ansatze. The variational imaginary time evolution closely approximates the exact imaginary time evolution described by Eq. (2).

In the minimal basis, the LiH molecule has 12 spin-orbitals, but the Hamiltonian can be reduced to act on six qubits, as described in Ref. [46] and the Supplementary Materials. We use a hardware efficient ansatz shown in

the Supplementary Materials for our simulation, with 42 parameters. We consider an internuclear distance of $R = 1.45\mathring{A}$, for which the exact ground state energy is $E_{\min} = -7.8807$ Hartree. We compare the performance of our variational imaginary time evolution method with that of gradient descent, with time step $\delta\tau = 0.01$. We applied both methods to 112 different trials, each initialised with different, random starting values of the parameters. As exemplified by one specific instance, shown in Fig. 3(b), we find that our method can almost always (with more than 90% probability) find the ground state energy, while gradient descent often becomes trapped in local minima (see Fig. 7 in the Supplementary Materials).

Discussion.— In this letter, we propose a method to efficiently simulate imaginary time evolution using a quantum computer. Our hybrid algorithm combines the ability of quantum computers to construct parametrised trial wavefunctions, with the ability of classical computers to update the parameter values. We apply our method to finding the ground state energy of quantum systems, and test its performance by finding the ground state energy of H_2 and LiH. We expect that our method would also be suitable for solving general optimisation problems, such as 3-SAT.

Our method can also be used to prepare a thermal (Gibbs) state, $\rho_T = e^{-H/T}/\mathrm{Tr}[e^{-H/T}]$ of Hamiltonian H at temperature T. Sampling from a Gibbs distribution is an important aspect of many machine learning algorithms, and so we believe that our method is applicable to problems in quantum machine learning. Moreover, while previous methods to prepare the Gibbs state [47, 48] require long gate sequences (and hence, fault tolerance), our method can be implemented using a shallow circuit. Our algorithm can also be combined with recently proposed error mitigation techniques [24, 49, 50], which protect shallow circuits from error accumulation, without the additional qubit resources required for full error correction. As a result, it is suitable for implementation on current and near-term quantum hardware.

Although exact imaginary time evolution deterministically propagates a good initial state to the ground state in the limit that $\tau \to \infty$, our variational method may still become trapped in local minima, if the chosen ansatz is not sufficiently powerful. In future works, we will investigate how our method may be optimally applied to a variety of tasks. This will include developing suitable ansatze for a range of problems. Another interesting avenue for further study is a more general imaginary time evolution $e^{-f(H,\tau)}$, where f is a function, such as $f(H,\tau) = (H-\lambda)^2\tau$. This may enable the investigation of excited states, or expedite convergence to the ground state.

This work was supported by BP plc and by the EPSRC National Quantum Technology Hub in Networked Quantum Information Technology (EP/M013243/1). YL is supported by NSAF (Grant No. U1730449). We acknowledge Tyson Jones for insightful discussions. SB and YL thank Sergey Bravyi for suggesting an investigation into the imaginary time variant of the algorithm in Ref. [24].

* simon.benjamin@materials.ox.ac.uk
† xiao.yuan.ph@gmail.com

[1] G. C. Wick, Phys. Rev. **96**, 1124 (1954).
[2] M. H. Poincaré, Rendiconti del Circolo Matematico di Palermo (1884-1940) **21**, 129 (1906).
[3] J. J. Sakurai and J. Napolitano, *Modern quantum mechanics* (Cambridge University Press, 2017).
[4] J. C. Baez and B. S. Pollard, Entropy **17**, 772 (2015).
[5] F. Verstraete, J. J. García-Ripoll, and J. I. Cirac, Phys. Rev. Lett. **93**, 207204 (2004).
[6] M. Zwolak and G. Vidal, Phys. Rev. Lett. **93**, 207205 (2004).
[7] F. A. Wolf, A. Go, I. P. McCulloch, A. J. Millis, and U. Schollwöck, Phys. Rev. X **5**, 041032 (2015).
[8] L. Lehtovaara, J. Toivanen, and J. Eloranta, Journal of Computational Physics **221**, 148 (2007).
[9] C. V. Kraus and J. I. Cirac, New Journal of Physics **12**, 113004 (2010).
[10] J. R. McClean and A. Aspuru-Guzik, RSC Adv. **5**, 102277 (2015).
[11] T. Shi, E. Demler, and J. I. Cirac, Annals of Physics **390**, 245 (2018).
[12] J. R. McClean, J. A. Parkhill, and A. Aspuru-Guzik, **110**, E3901 (2013).
[13] J. R. McClean and A. Aspuru-Guzik, Phys. Rev. A **91**, 012311 (2015).
[14] R. P. Feynman, International Journal of Theoretical Physics **21**, 467 (1982).
[15] F. Dalfovo, S. Giorgini, L. P. Pitaevskii, and S. Stringari, Rev. Mod. Phys. **71**, 463 (1999).
[16] I. M. Georgescu, S. Ashhab, and F. Nori, Rev. Mod. Phys. **86**, 153 (2014).
[17] M. A. Nielsen and I. Chuang, "Quantum computation and quantum information," (2002).
[18] H. F. Trotter, Proceedings of the American Mathematical Society **10**, 545 (1959).
[19] D. S. Abrams and S. Lloyd, Phys. Rev. Lett. **79**, 2586 (1997).
[20] R. Jackiw and A. Kerman, Physics Letters A **71**, 158 (1979).
[21] P. Kramer, Journal of Physics: Conference Series **99**, 012009.
[22] J. Haegeman, J. I. Cirac, T. J. Osborne, I. Pižorn, H. Verschelde, and F. Verstraete, Phys. Rev. Lett. **107**, 070601 (2011).
[23] Y. Ashida, T. Shi, M. C. Bañuls, J. I. Cirac, and E. Demler, arXiv preprint arXiv:1802.03861 (2018).
[24] Y. Li and S. C. Benjamin, Phys. Rev. X **7**, 021050 (2017).
[25] A. Peruzzo, J. McClean, P. Shadbolt, M.-H. Yung, X.-Q. Zhou, P. J. Love, A. Aspuru-Guzik, and J. L. Obrien, Nature communications **5** (2014).
[26] Y. Wang, F. Dolde, J. Biamonte, R. Babbush, V. Bergholm, S. Yang, I. Jakobi, P. Neumann,

A. Aspuru-Guzik, J. D. Whitfield, *et al.*, ACS nano **9**, 7769 (2015).
[27] P. J. J. O'Malley, R. Babbush, I. D. Kivlichan, J. Romero, J. R. McClean, R. Barends, J. Kelly, P. Roushan, A. Tranter, N. Ding, B. Campbell, Y. Chen, Z. Chen, B. Chiaro, A. Dunsworth, A. G. Fowler, E. Jeffrey, E. Lucero, A. Megrant, J. Y. Mutus, M. Neeley, C. Neill, C. Quintana, D. Sank, A. Vainsencher, J. Wenner, T. C. White, P. V. Coveney, P. J. Love, H. Neven, A. Aspuru-Guzik, and J. M. Martinis, Phys. Rev. X **6**, 031007 (2016).
[28] Y. Shen, X. Zhang, S. Zhang, J.-N. Zhang, M.-H. Yung, and K. Kim, Phys. Rev. A **95**, 020501 (2017).
[29] J. R. McClean, J. Romero, R. Babbush, and A. Aspuru-Guzik, New Journal of Physics **18**, 023023 (2016).
[30] S. Paesani, A. A. Gentile, R. Santagati, J. Wang, N. Wiebe, D. P. Tew, J. L. O'Brien, and M. G. Thompson, Phys. Rev. Lett. **118**, 100503 (2017).
[31] A. Kandala, A. Mezzacapo, K. Temme, M. Takita, M. Brink, J. M. Chow, and J. M. Gambetta, Nature **549**, 242 (2017).
[32] E. Farhi, J. Goldstone, and S. Gutmann, arXiv preprint arXiv:1411.4028 (2014).
[33] J. Preskill, arXiv preprint arXiv:1801.00862 (2018).
[34] A. Aspuru-Guzik, A. D. Dutoi, P. J. Love, and M. Head-Gordon, Science **309**, 1704 (2005).
[35] D. Poulin, A. Qarry, R. Somma, and F. Verstraete, Phys. Rev. Lett. **106**, 170501 (2011).
[36] F. Verstraete, V. Murg, and J. I. Cirac, Advances in Physics **57**, 143 (2008).
[37] K. B. Whaley, A. R. Dinner, and S. A. Rice, *Quantum information and computation for chemistry* (John Wiley & Sons, 2014).
[38] A. McLachlan, Molecular Physics **8**, 39 (1964).
[39] J. Romero, R. Babbush, J. R. McClean, C. Hempel, P. Love, and A. Aspuru-Guzik, arXiv preprint arXiv:1701.02691 (2017).
[40] P.-L. Dallaire-Demers, J. Romero, L. Veis, S. Sim, and A. Aspuru-Guzik, arXiv preprint arXiv:1801.01053 (2018).
[41] D. Wecker, M. B. Hastings, and M. Troyer, Phys. Rev. A **92**, 042303 (2015).
[42] S. B. Bravyi and A. Y. Kitaev, Annals of Physics **298**, 210 (2002).
[43] I. Kassal, J. D. Whitfield, A. Perdomo-Ortiz, M.-H. Yung, and A. Aspuru-Guzik, Annual review of physical chemistry **62**, 185 (2011).
[44] J. Olson, Y. Cao, J. Romero, P. Johnson, P.-L. Dallaire-Demers, N. Sawaya, P. Narang, I. Kivlichan, M. Wasielewski, and A. Aspuru-Guzik, ArXiv e-prints (2017), arXiv:1706.05413 [quant-ph].
[45] S. McArdle, S. Endo, S. Benjamin, and X. Yuan, in preparation.
[46] C. Hempel, C. Maier, J. Romero, J. McClean, T. Monz, H. Shen, P. Jurcevic, B. Lanyon, P. Love, R. Babbush, A. Aspuru-Guzik, R. Blatt, and C. Roos, ArXiv e-prints (2018), arXiv:1803.10238 [quant-ph].
[47] K. Temme, T. J. Osborne, K. G. Vollbrecht, D. Poulin, and F. Verstraete, Nature **471**, 87 (2011).
[48] A. Riera, C. Gogolin, and J. Eisert, Phys. Rev. Lett. **108**, 080402 (2012).
[49] K. Temme, S. Bravyi, and J. M. Gambetta, Phys. Rev. Lett. **119**, 180509 (2017).
[50] S. Endo, S. C. Benjamin, and Y. Li, arXiv preprint arXiv:1712.09271 (2017).
[51] J. T. Seeley, M. J. Richard, and P. J. Love, The Journal of Chemical Physics **137**, 224109 (2012), https://doi.org/10.1063/1.4768229.
[52] J. R. McClean, I. D. Kivlichan, D. S. Steiger, Y. Cao, E. S. Fried, C. Gidney, T. Häner, V. Havlíček, Z. Jiang, M. Neeley, *et al.*, arXiv preprint arXiv:1710.07629 (2017).
[53] S. Bravyi, J. M. Gambetta, A. Mezzacapo, and K. Temme, arXiv preprint arXiv:1701.08213 (2017).

Supplementary Materials

Variational simulation of imaginary time evolution

McLachlan's variational principle [38], applied to imaginary time evolution, is given by

$$\delta \| (\partial/\partial\tau + H - E_\tau) \left|\psi(\tau)\right\rangle \| = 0 \tag{12}$$

where

$$\| (\partial/\partial\tau + H - E_\tau) \left|\psi(\tau)\right\rangle \| = \left((\partial/\partial\tau + H - E_\tau) \left|\psi(\tau)\right\rangle\right)^\dagger (\partial/\partial\tau + H - E_\tau) \left|\psi(\tau)\right\rangle, \tag{13}$$

and $E_\tau = \langle\psi(\tau)|H|\psi(\tau)\rangle$. For a general quantum state, McLachlan's variational principle recovers the imaginary time evolution

$$\frac{\partial \left|\psi(\tau)\right\rangle}{\partial\tau} = -(H - E_\tau) \left|\psi(\tau)\right\rangle. \tag{14}$$

If we consider a subspace of the whole Hilbert space, which can be reached using the ansatz $|\phi(\tau)\rangle = |\phi(\theta_1, \theta_2, \ldots, \theta_N)\rangle$, we can project the imaginary time evolution onto the subspace using McLachlan's variational principle. Replacing $|\psi(\tau)\rangle$ with $|\phi(\tau)\rangle$, yields

$$\begin{aligned}
\| (\partial/\partial\tau + H - E_\tau) \left|\phi(\tau)\right\rangle \| &= \left((\partial/\partial\tau + H - E_\tau) \left|\psi(\tau)\right\rangle\right)^\dagger (\partial/\partial\tau + H - E_\tau) \left|\psi(\tau)\right\rangle, \\
&= \sum_{i,j} \frac{\partial \left\langle\phi(\tau)\right|}{\partial\theta_i} \frac{\partial \left|\phi(\tau)\right\rangle}{\partial\theta_j} \dot\theta_i \dot\theta_j + \sum_i \frac{\partial \left\langle\phi(\tau)\right|}{\partial\theta_i} (H - E_\tau) \left|\phi(\tau)\right\rangle \dot\theta_i \\
&+ \sum_i \left\langle\phi(\tau)\right| (H - E_\tau) \frac{\partial \left|\phi(\tau)\right\rangle}{\partial\theta_i} \dot\theta_i + \left\langle\phi(\tau)\right| (H - E_\tau)^2 \left|\phi(\tau)\right\rangle.
\end{aligned} \tag{15}$$

Focusing on $\dot\theta_i$, we obtain

$$\begin{aligned}
\frac{\partial \| (\partial/\partial\tau + H - E_\tau) \left|\phi(\tau)\right\rangle \|}{\partial\dot\theta_i} &= \sum_j \left(\frac{\partial \left\langle\phi(\tau)\right|}{\partial\theta_i} \frac{\partial \left|\phi(\tau)\right\rangle}{\partial\theta_j} + \frac{\partial \left\langle\phi(\tau)\right|}{\partial\theta_j} \frac{\partial \left|\phi(\tau)\right\rangle}{\partial\theta_i} \right) \dot\theta_j \\
&+ \frac{\partial \left\langle\phi(\tau)\right|}{\partial\theta_i} (H - E_\tau) \left|\phi(\tau)\right\rangle + \left\langle\phi(\tau)\right| (H - E_\tau) \frac{\partial \left|\phi(\tau)\right\rangle}{\partial\theta_i}.
\end{aligned} \tag{16}$$

Considering the normalisation condition for the trial state $|\phi(\tau)\rangle$,

$$\left\langle\phi(\tau)\right| \left|\phi(\tau)\right\rangle = 1, \tag{17}$$

we have

$$E_\tau \frac{\partial \langle\phi(\tau)|\phi(\tau)\rangle}{\partial\theta_i} = E_\tau \left(\frac{\partial \left\langle\phi(\tau)\right|}{\partial\theta_i} \left|\phi(\tau)\right\rangle + \left\langle\phi(\tau)\right| \frac{\partial \left|\phi(\tau)\right\rangle}{\partial\theta_i} \right) = 0, \tag{18}$$

and the derivative is simplified to

$$\frac{\partial \| (\partial/\partial\tau + H - E_\tau) \left|\phi(\tau)\right\rangle \|}{\partial\dot\theta_i} = \sum_j A_{ij} \dot\theta_j - C_i. \tag{19}$$

where

$$\begin{aligned}
A_{ij} &= \Re \left(\frac{\partial \left\langle\phi(\tau)\right|}{\partial\theta_i} \frac{\partial \left|\phi(\tau)\right\rangle}{\partial\theta_j} \right), \\
C_i &= -\Re \left(\frac{\partial \left\langle\phi(\tau)\right|}{\partial\theta_i} H \left|\phi(\tau)\right\rangle \right).
\end{aligned} \tag{20}$$

McLachlan's variational principle requires

$$\frac{\partial \| (\partial/\partial\tau + H - E_\tau) \left|\phi(\tau)\right\rangle \|}{\partial\dot\theta_j} = 0, \tag{21}$$

which is equivalent to the differential equation of the parameters

$$\sum_j A_{ij}\dot{\theta}_j = C_i. \tag{22}$$

Denoting $E(\tau) = \langle\phi(\tau)| H |\phi(\tau)\rangle$, we can show that the average energy always decreases by following our imaginary time evolution algorithm, for a sufficiently small stepsize;

$$\begin{aligned}
\frac{dE(\tau)}{d\tau} &= \Re\left(\langle\phi(\tau)| H \frac{d|\phi(\tau)\rangle}{d\tau}\right), \\
&= \sum_i \Re\left(\langle\phi(\tau)| H \frac{\partial|\phi(\tau)\rangle}{\partial\theta_i}\dot{\theta}_i\right), \\
&= -\sum_i C_i\dot{\theta}_i, \\
&= -\sum_i C_i A^{-1}_{ij} C_j, \\
&\leq 0.
\end{aligned} \tag{23}$$

The third line follows from the definition of C_i; the fourth line follows from the differential equation of $\dot{\theta}$; the last line is true when A^{-1} is positive. First, we show matrix A is positive. We consider an arbitrary vector $x = (x_1, x_2, \ldots, x_N)^T$, and calculate $x^\dagger \cdot A \cdot x$,

$$\begin{aligned}
x^\dagger \cdot A \cdot x &= \sum_{i,j} x_i^* A_{ij} x_j, \\
&= \sum_{i,j} x_i^* \Re\left(\frac{\partial\langle\phi(\tau)|}{\partial\theta_i}\frac{\partial|\phi(\tau)\rangle}{\partial\theta_j}\right) x_j, \\
&= \sum_{i,j} x_i^* \frac{\partial\langle\phi(\tau)|}{\partial\theta_i}\frac{\partial|\phi(\tau)\rangle}{\partial\theta_j} x_j + \sum_{i,j} x_i^* \frac{\partial\langle\phi(\tau)|}{\partial\theta_j}\frac{\partial|\phi(\tau)\rangle}{\partial\theta_i} x_j,
\end{aligned} \tag{24}$$

Denote $|\Phi\rangle = \sum_i x_i \frac{\partial|\phi(\tau)\rangle}{\partial\theta_i}$, then the first term equals

$$\sum_{i,j} x_i^* \frac{\partial\langle\phi(\tau)|}{\partial\theta_i}\frac{\partial|\phi(\tau)\rangle}{\partial\theta_j} x_j = \langle\Phi|\Phi\rangle \geq 0. \tag{25}$$

Similarly, we can show that the second term is also nonnegative. Therefore, $x^\dagger \cdot A \cdot x \geq 0, \forall x$ and A is nonnegative. In practice, when A has eigenvalues with value zero, A is not invertible. However, in our simulation, we define the inverse of A to be only the inverse of the nonnegative eigenvalues. Suppose U is the transformation that diagonalises A, i.e., $G_{i,j} = (UAU^\dagger)_{i,j} = 0, \forall i \neq j$. Then, we define G^{-1} by

$$G^{-1}_{i,j} = \begin{cases} \frac{1}{G_{i,j}} & i = j, G_{i,j} \neq 0, \\ 0 & i = j, G_{i,j} = 0, \\ 0 & i \neq j. \end{cases} \tag{26}$$

The inverse of A is thus defined by

$$A^{-1} = U^\dagger G^{-1} U. \tag{27}$$

Because A has nonnegative eigenvalues, G, G^{-1}, and hence A^{-1} all have nonnegative eigenvalues.

Evaluating A and C with quantum circuits

In this section, we review the quantum circuit that can efficiently evaluate the coefficients A and C introduced in Ref. [24, 39, 40].

Without loss of generality, we can assume that each unitary gate $U_i(\theta_i)$ in our circuit depends only on parameter θ_i (since multiple parameter gates can be decomposed into this form). Suppose each U_i is a rotation or a controlled rotation gate, its derivative can be expressed by

$$\frac{\partial U_i(\theta_i)}{\partial \theta_i} = \sum_k f_{k,i} U_i(\theta_i) \sigma_{k,i}, \tag{28}$$

with unitary operator $\sigma_{k,i}$ and scalar parameters $f_{k,i}$. The derivative of the trial state is

$$\frac{\partial \left|\phi(\tau)\right\rangle}{\partial \theta_i} = \sum_k f_{k,i} \tilde{V}_{k,i} \left|\bar{0}\right\rangle, \tag{29}$$

with

$$\tilde{V}_{k,i} = U_N(\theta_N) \dots U_{i+1}(\theta_{i+1}) U_i(\theta_i) \sigma_{k,i} \dots U_2(\theta_2) U_1(\theta_1). \tag{30}$$

In practice, there are only one or two terms, $f_{k,i}\sigma_{k,i}$, for each derivative. For example, when $U_i(\theta_i)$ is a single qubit rotation $R^Z_{\theta_i} = e^{-i\theta_i \sigma_Z/2}$, the derivative $\partial U_i(\theta_i)/\partial\theta_i = -i/2 \times Z e^{-i\theta_i Z/2}$, and the derivative of the trial state $\partial \left|\phi(\tau)\right\rangle/\partial\theta_i$ can be prepared by adding an extra Z gate with a constant factor $-i/2$. When $U_i(\theta_i)$ is a control rotation such as $\left|0\right\rangle\left\langle 0\right| \otimes I + \left|1\right\rangle\left\langle 1\right| \otimes R^Z_{\theta_i}$, the derivative $\partial U_i(\theta_i)/\partial\theta_i = \left|1\right\rangle\left\langle 1\right| \otimes \partial R^Z_{\theta_i}/\partial\theta_i = -i/2 \times \left|1\right\rangle\left\langle 1\right| \otimes Z e^{-i\theta_i Z/2}$. By choosing $\sigma_{1,i} = I \otimes Z$, $\sigma_{1,i} = Z \otimes Z$, $f_{1,i} = -i/4$, and $f_{2,i} = i/4$, we can show Eq. (28).

Therefore, the coefficients $A_{i,j}$ and C_i are given by

$$\begin{aligned} A_{i,j} &= \Re\left(\sum_{k,l} f^*_{k,i} f_{l,j} \left\langle \bar{0}\right| \tilde{V}^\dagger_{k,i} \tilde{V}_{l,j} \left|\bar{0}\right\rangle \right), \\ C_i &= \Re\left(\sum_{k,l} f^*_{k,i} \lambda_l \left\langle \bar{0}\right| \tilde{V}^\dagger_{k,i} h_l V \left|\bar{0}\right\rangle \right). \end{aligned} \tag{31}$$

All the terms of the summation follow the general form $a\Re(e^{i\theta}\left\langle \bar{0}\right| U \left|\bar{0}\right\rangle)$ and can be evaluated by the circuit in Fig. 1 in the main text.

In practice, we do not need to realize the whole controlled-U gate and can instead use a much easier circuit. For example, for the term $\Re(f^*_{k,i} f_{l,j} \left\langle \bar{0}\right| \tilde{V}^\dagger_{k,i} \tilde{V}_{l,j} \left|\bar{0}\right\rangle)$, we can let $f^*_{k,i} f_{l,j} = a e^{i\theta}$ and

$$\left\langle \bar{0}\right| \tilde{V}^\dagger_{k,i} \tilde{V}_{l,j} \left|\bar{0}\right\rangle = \left\langle \bar{0}\right| U_1^\dagger \dots U_{i-1}^\dagger \sigma^\dagger_{k,i} U_i^\dagger \dots U_N^\dagger U_N \dots U_j \sigma_{l,j} U_{j-1} \dots U_1 \left|\bar{0}\right\rangle. \tag{32}$$

Suppose $i < j$, then

$$\left\langle \bar{0}\right| \tilde{V}^\dagger_{k,i} \tilde{V}_{l,j} \left|\bar{0}\right\rangle = \left\langle \bar{0}\right| U_1^\dagger \dots U_{i-1}^\dagger \sigma^\dagger_{k,i} U_i^\dagger \dots U_{j-1}^\dagger \sigma_{l,j} U_{j-1} \dots U_i \dots U_1 \left|\bar{0}\right\rangle, \tag{33}$$

and $\Re(e^{i\theta}\left\langle \bar{0}\right| \tilde{V}^\dagger_{k,i} \tilde{V}_{l,j} \left|\bar{0}\right\rangle)$ can be measured by the circuit in Fig. 4. The terms for C can be measured similarly.

Computational chemistry background

One of the central problems in computational chemistry is finding the ground state energy of molecules. This calculation is classically intractable, due to the exponential growth of Hilbert space with the number of electrons in the molecule. However, it has been shown that quantum computers are able to solve this problem efficiently [34]. The Hamiltonian of a molecule consisting of M nuclei (of mass M_I, position $\mathbf{R}_I$, and charge Z_I) and N electrons (with position $\mathbf{r}_i$) is

$$H = -\sum_i \frac{\hbar^2}{2m_e}\nabla_i^2 - \sum_I \frac{\hbar^2}{2M_I}\nabla_I^2 - \sum_{i,I} \frac{e^2}{4\pi\epsilon_0}\frac{Z_I}{|\mathbf{r}_i - \mathbf{R}_I|} + \frac{1}{2}\sum_{i\neq j} \frac{e^2}{4\pi\epsilon_0}\frac{1}{|\mathbf{r}_i - \mathbf{r}_j|} + \frac{1}{2}\sum_{I\neq J} \frac{e^2}{4\pi\epsilon_0}\frac{Z_I Z_J}{|\mathbf{R}_I - \mathbf{R}_J|}. \tag{34}$$

Because the nuclei are orders of magnitude more massive than the electrons, we apply the Born-Oppenheimer approximation, and treat the nuclei as classical, fixed point charges. After this approximation, the eigenvalue equation

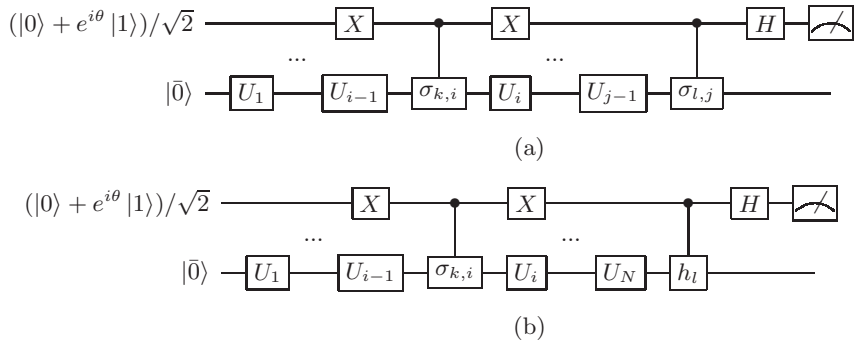

FIG. 4. Quantum circuits that evaluate (a) $\Re(e^{i\theta}\langle\bar{0}|\tilde{V}_{k,i}^{\dagger}\tilde{V}_{l,j}|\bar{0}\rangle)$ and (b) $\Re(e^{i\theta}\langle\bar{0}|\tilde{V}_{k,i}^{\dagger}h_l V|\bar{0}\rangle)$. When $\sigma_{k,i}$ is Hermitian, the X gates acting on the ancilla qubit can be also omitted.

we seek to solve (in atomic units) is given by

$$\left[-\sum_i \frac{\nabla_i^2}{2} - \sum_{i,I}\frac{Z_I}{|\mathbf{r}_i - \mathbf{R}_I|} + \frac{1}{2}\sum_{i\neq j}\frac{1}{|\mathbf{r}_i - \mathbf{r}_j|}\right]|\psi\rangle = E|\psi\rangle\,, \tag{35}$$

where $|\psi\rangle$ is an energy eigenstate of the Hamiltonian, with energy eigenvalue E.

To solve this problem using a quantum computer, we first transform it into the second quantised form. We project the Hamiltonian onto a finite number of basis wave functions, $\{\phi_p\}$, which approximate spin orbitals. Electrons are excited into, or de-excited out of, these orbitals by fermionic creation ($a_p^{\dagger}$) or annihilation (a_p) operators, respectively. These operators obey fermionic anti-commutation relations, which enforces the antisymmetry of the wavefunction, a consequence of the Pauli exclusion principle. In the second quantised representation, the electronic Hamiltonian is written as

$$H = \sum_{p,q} h_{pq} a_p^{\dagger} a_q + \frac{1}{2}\sum_{p,q,r,s} h_{pqrs} a_p^{\dagger} a_q^{\dagger} a_r a_s, \tag{36}$$

with

$$\begin{aligned} h_{pq} &= \int d\mathbf{x}\phi_p^*(\mathbf{x})\left(\frac{\nabla_i^2}{2} - \sum_I \frac{Z_I}{|\mathbf{r}-\mathbf{R}_I|}\right)\phi_q(\mathbf{x}), \\ h_{pqrs} &= \int d\mathbf{x}_1 d\mathbf{x}_2 \frac{\phi_p^*(\mathbf{x}_1)\phi_q^*(\mathbf{x}_2)\phi_s(\mathbf{x}_1)\phi_r(\mathbf{x}_2)}{|\mathbf{r}_1 - \mathbf{r}_2|}, \end{aligned} \tag{37}$$

where $\mathbf{x}$ is a spatial and spin coordinate. This Hamiltonian in general contains N_{SO}^4 terms, where N_{SO} is the number of orbitals considered. This fermionic Hamiltonian must then be transformed into a Hamiltonian acting on qubits. This is achieved using the Jordan-Wigner, or Bravyi-Kitaev transformations, which are described in Ref. [51].

Hydrogen

In our simulations, we consider the Hydrogen molecule in the minimal STO-6G basis. This means that only the minimum number of orbitals to describe the electrons are considered. 'STO-nG' means that a linear combination of n Gaussian functions are used to approximate a Slater-type-orbital, which describes the electron wavefunction. Each Hydrogen atom contributes a single $1S$ orbital. As a result of spin multiplicity, there are four spin orbitals in total.

We are able to construct the molecular orbitals for H_2 by manually (anti)symmetrising the orbitals. These are

$$\begin{aligned}
|\phi_0\rangle = |\sigma_{g\uparrow}\rangle &= \frac{1}{\sqrt{2}}(|1S_{1\uparrow}\rangle + |1S_{2\uparrow}\rangle), \\
|\phi_1\rangle = |\sigma_{g\downarrow}\rangle &= \frac{1}{\sqrt{2}}(|1S_{1\downarrow}\rangle + |1S_{2\downarrow}\rangle), \\
|\phi_2\rangle = |\sigma_{u\uparrow}\rangle &= \frac{1}{\sqrt{2}}(|1S_{1\uparrow}\rangle - |1S_{2\uparrow}\rangle), \\
|\phi_3\rangle = |\sigma_{u\downarrow}\rangle &= \frac{1}{\sqrt{2}}(|1S_{1\downarrow}\rangle - |1S_{2\downarrow}\rangle),
\end{aligned} \tag{38}$$

where the subscripts on the $1S$ orbitals denote the spin of the electron in that orbital, and which of the two hydrogen atoms the orbital is centred on. By following the procedure in Ref. [51], the qubit Hamiltonian for H_2 in the BK representation can be obtained. This 4 qubit Hamiltonian is given by

$$\begin{aligned}
H = h_0 I + h_1 Z_0 + h_2 Z_1 + h_3 Z_2 + h_4 Z_0 Z_1 + h_5 Z_0 Z_2 + h_6 Z_1 Z_3 + h_7 X_0 Z_1 X_2 + h_8 Y_0 Z_1 Y_2 \\
+ h_9 Z_0 Z_1 Z_2 + h_{10} Z_0 Z_2 Z_3 + h_{11} Z_1 Z_2 Z_3 + h_{12} X_0 Z_1 X_2 Z_3 + h_{13} Y_0 Z_1 Y_2 Z_3 + h_{14} Z_0 Z_1 Z_2 Z_3.
\end{aligned} \tag{39}$$

As this Hamiltonian only acts off diagonally on qubits 0 and 2 [27, 31], it can be reduced to

$$H = g_0 + g_1 Z_0 + g_2 Z_1 + g_3 Z_0 Z_1 + g_4 Y_0 Y_1 + g_5 X_0 X_1, \tag{40}$$

which only acts on two qubits.

Lithium Hydride

In our simulations, we consider Lithium Hydride in the minimal STO-3G basis. The Lithium atom has 3 electrons, and so contributes a $1S, 2S, 2P_x, 2P_y$ and $2P_z$ orbital to the basis, while the Hydrogen atom contributes a single $1S$ orbital. With spin multiplicity, this makes 12 orbitals in total. However, we are able to reduce the number of spin orbitals required by considering their expected occupation. This reduces the qubit resources required for our calculation. In computational chemistry, the subset of spin orbitals included in a calculation is called the active space.

We first obtain the one electron reduced density matrix (1-RDM) for LiH, using a classically tractable CISD (configuration interaction, single and double excitations) calculation. The 1-RDM for a distance of 1.45 $\mathring{A}$ is shown below,

$$\begin{pmatrix}
0.99996 & 0 & 0.00024 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0.99996 & 0 & 0.00024 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0.00024 & 0 & 0.97988 & 0 & 0.0334 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0.00024 & 0 & 0.97988 & 0 & 0.0334 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0.0334 & 0 & 0.00483 & 0 & 0 & 0 & 0 & 0 & 0.00691 & 0 \\
0 & 0 & 0 & 0.0334 & 0 & 0.00483 & 0 & 0 & 0 & 0 & 0 & 0.00691 \\
0 & 0 & 0 & 0 & 0 & 0 & 0.00086 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0.00086 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0.00086 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0.00086 & 0 & 0 \\
0 & 0 & 0 & 0 & 0.00691 & 0 & 0 & 0 & 0 & 0 & 0.01363 & 0 \\
0 & 0 & 0 & 0 & 0 & 0.00691 & 0 & 0 & 0 & 0 & 0 & 0.01363
\end{pmatrix}. \tag{41}$$

The diagonal elements of the 1-RDM are the occupation numbers of the corresponding canonical orbitals (the orbitals contributed by the individual atoms, eg. $1S_{Li}$). In order to reduce our active space, we first perform a unitary rotation

of the 1-RDM, such that it becomes a diagonal matrix,

$$\begin{pmatrix} 0.99996 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0.99996 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0.98103 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0.98103 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0.01727 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0.01727 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0.00086 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0.00086 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0.00086 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0.00086 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0.00000 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0.00000 \end{pmatrix}. \tag{42}$$

This gives the 1-RDM in terms of natural molecular orbitals (NMOs). The diagonal entries are called the natural orbital occupation numbers (NOONs). The Hamiltonian of LiH must also be rotated, using the same unitary matrix used to diagonalise the 1-RDM. This is equivalent to performing a change of basis, from the canonical orbital basis to the natural molecular orbital basis.

As can be seen, the first and the second orbitals have a NOON close to unity, and so are very likely to be occupied. As a result, we 'freeze' these core orbitals, and consider them to always be filled. We can then remove any terms containing $a_0^\dagger, a_0, a_1^\dagger, a_1$ from the LiH fermionic Hamiltonian. We also notice that the last two orbitals have NOONs close to zero. As a result, we assume that these orbitals are never occupied, and so remove them from the Hamiltonian. This leaves a fermionic Hamiltonian acting on 8 spin orbitals. We then map this fermionic Hamiltonian to a qubit Hamiltonian, using the BK transformation. All of these steps were carried out using OpenFermion [52], an electronic structure package to transform computational chemistry problems into a form that is suitable for investigation using a quantum computer. The BK transformation enables us to further reduce the Hamiltonian by two qubits, using the spin and electron number symmetries of the molecule [31, 53]. This results in a Hamiltonian for LiH which acts on 6 qubits. A similar encoding and reduction procedure for LiH is also described in Ref. [46].

Hardware-efficient ansatz

A hardware-efficient ansatz was recently realised using superconducting qubits [31]. Following a similar, but slightly different structure, the general hardware-efficient ansatz we considered in this work is shown in Fig. 5 and the ansatz for simulating LiH is shown in Fig. 6.

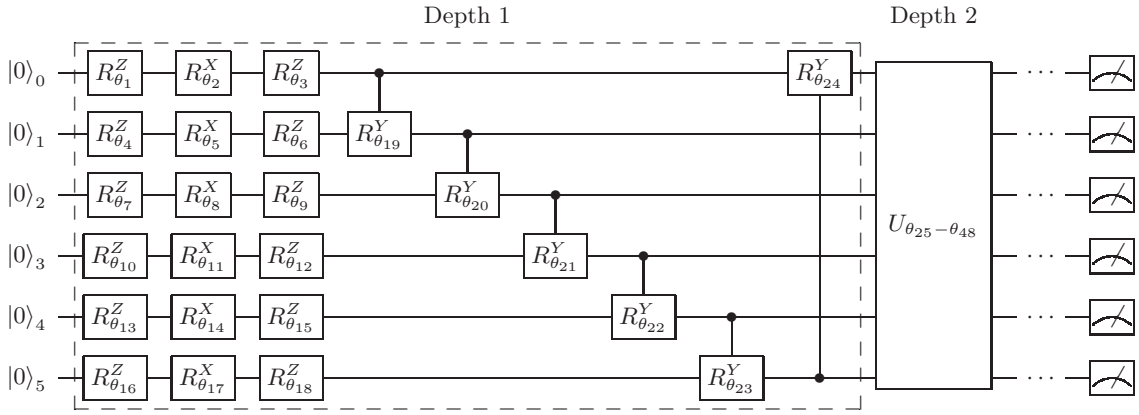

FIG. 5. General hardware efficient ansatz. We repeat the circuit structure of the first block to depth M. In the first block, we set the parameters of the first Z rotation gates to be 0.

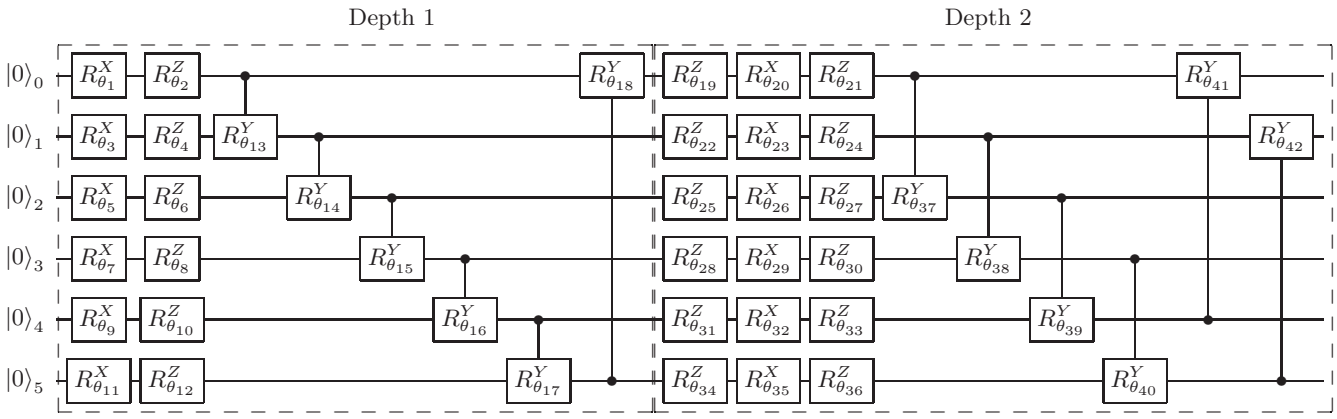

FIG. 6. The hardware efficient ansatz for the LiH simulation results presented in this paper. In our simulation, we used this two block circuit for the LiH molecule. In total, there are 42 parameters.

Additional LiH simulation results

The exact ground state energy of LiH, at an internuclear distance of $R = 1.45\mathring{A}$, is $E_{\min} = -7.8807$ Hartree (when working in the STO-3G basis). We compare the performance of our variational imaginary time evolution method with that of gradient descent, with time step $\delta\tau = 0.01$. We applied both methods to 112 different trials, each initialised with different, random starting values of the parameters. As shown in Fig. 7, our method can almost always (with more than 90% probability) find the ground state energy, while gradient descent often becomes trapped in local minima.

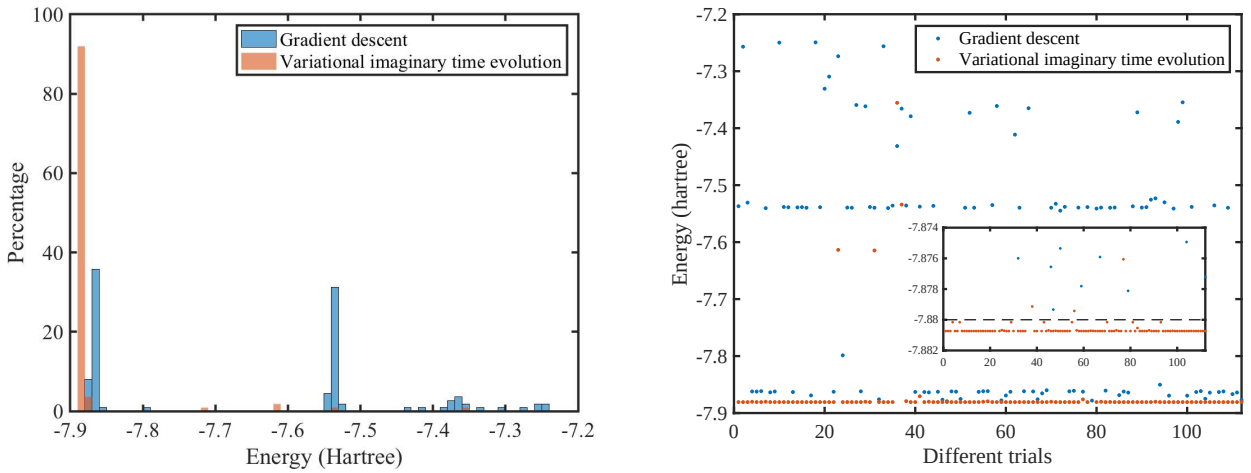

FIG. 7. (Color online) Simulation results of the 112 trials with different, random starting values of the parameters. (a) Histogram of the 112 trials. The variational imaginary time evolution reaches within Chemical Accuracy (10^{-3} Hartree) of the true ground state $E_0 = -7.8807$ Hartree in 103 out of 112 trials. In contrast, gradient descent generally gets trapped in local minima. (b) Scatter graph of the 112 trials. The dashed line in the inset plot marks results which fall within Chemical Accuracy of the ground state energy.