Lecture 11: Variational imaginary time evolution
Reading: McArdle et al, 2018.

1 Introduction

We now move to other approaches to obtain ground states of Hamiltonians as well as thermal averages. The first one we will consider is still a variational scheme, but one that avoids a classical non-linear optimization as in the standard VQE approach. It does so by using a QC to evolve a state in imaginary time in steps, where the QC is used to obtain matrix elements for a linear system, the solution of which provides the variational parameters at the next imaginary time step. By sequentially repeating this process for steps in imaginary time, we will eventually arrive at the ground state (or at a thermal average if we terminate the time stepping at the right time).

Say we have a Hamiltonian $H = \sum_i \lambda_i h_i$ with the coefficients λ_i real and each h_i a string of Pauli operators as usual. If the system register starts in an arbitrary state $|\psi\rangle$, the normalized imaginary time evolution (ITE) produces the state:

$$|\psi(\beta)\rangle = A(\beta)e^{-H\beta}\,|\psi(0)\rangle \tag{1.1}$$

where the normalization constant $A(\beta) = \langle\psi(0)|\exp(-2H\beta)|\psi(0)\rangle^{-1/2}$.

If the initial state has overlap with the ground state, as $\beta \to \infty$, $|\psi(\beta)\rangle \to |g\rangle$. To reach the ground state, we need to be able to evolve a state in imaginary time. The IT Schrodinger equation is given by:

$$\frac{\partial\,|\psi(\beta)\rangle}{\partial\beta} = -(H - E_\beta)\,|\psi(\beta)\rangle \tag{1.2}$$

where $E_\beta = \langle\psi(\beta)|H|\psi(\beta)\rangle$ is introduced for normalization. In the present variational scheme, we use a variational ansatz for the wavefunction as $|\psi(\beta)\rangle = |\phi(\vec{\theta}(\beta))\rangle$ where $\vec{\theta}(\beta)$ is a vector of parameters just as in VQE. We prepare the state $|\phi\rangle$ using a sequence of unitaries:

$$|\phi\rangle = U_N(\theta_N)...U_1(\theta_1)\,|\bar{0}\rangle = V(\vec{\theta})\,|\bar{0}\rangle \tag{1.3}$$

We now want to simulate ITE of this trial state. We do so by formulating a variational principle:

$$\delta||(\partial/\partial\beta + H - E)\,|\psi(\beta)\rangle\,||= 0 \tag{1.4}$$

where δ indicates a variation of the functional. If this variational principle is satisfied, it implies that the original IT Schrodinger equation is also satisfied.

Now we project this equation onto the ansatz wavefunction. The equations are messy and are given in the referenced paper (and I will give them in class). The main point is that out of all of this derivation comes a linear system that must be solved for derivatives of the parameters with respect to imaginary time, $\dot{\theta}$:

$$A_{ij}\dot{\theta}_j = c_i \tag{1.5}$$

where the matrix elements for the matrix and right-hand side look like:

$$A_{ij} = Re\left(\frac{\partial \langle\phi|}{\partial\theta_i}\frac{\partial |\phi\rangle}{\partial\theta_j}\right) \tag{1.6}$$

$$c_i = -Re\left(\frac{\partial \langle\phi|}{\partial\theta_i}H|\phi\rangle\right) \tag{1.7}$$

In the SI of the paper, it is shown that E_β is monotonically non-increasing, indicating that ITE always tends to the best ground state energy within the constraints of the ansatz.

2 Getting the matrix elements

Now the question is, how to get the matrix entries, which are generally involve gradient terms like $\partial|\phi\rangle/\partial\theta_i$. Considering that by definition $|\phi\rangle = U_N(\theta_N)...U_1(\theta_1)|0\rangle$, the derivative term for θ_i is:

$$\frac{\partial U_i}{\partial\theta_i} = \sum_k f_{ki}U_i\sigma_{ki} \tag{2.1}$$

where the assumption is that U_i is a rotation or controlled-rotation gate. Here σ_k is some Pauli operator string and f_k are scalar parameters. For a 1-qubit rotation gate, the sum would consist of just one term. As an example, consider a Z rotation gate $U_i = \exp(-i\sigma_z\theta_i/2)$.

$$\frac{\partial U_i}{\partial\theta_i} = -\frac{i}{2}\sigma_z U_i \tag{2.2}$$

Therefore, the overall gradient term is:

$$\frac{\partial|\phi\rangle}{\partial\theta_i} = \sum_k f_{k,i}\tilde{V}_{k,i}|0\rangle \tag{2.3}$$

$$\tilde{V}_{k,i} = U_N(\theta_N)...U_{i+1}U_i\sigma_{k,i}...U_1 \tag{2.4}$$

Therefore, the matrix elements are of the form:

$$A_{ij} = Re\left(\sum_{k,l} f^*_{k,i}f_{l,j}\langle 0|\tilde{V}^\dagger_{k,i}\tilde{V}_{l,j}|0\rangle\right) \tag{2.5}$$

$$c_i = Re\left(\sum_{k,\alpha} f^*_{k,i}\lambda_\alpha\langle 0|\tilde{V}^\dagger_{k,i}h_\alpha V|0\rangle\right) \tag{2.6}$$

These terms are generally proportional to $Re(\exp(i\theta)\langle 0|U|0\rangle)$. How do we get such a term evaluated on a quantum computer?

The circuit to do it was presented in a few papers in the early 2000s [1–3]. We will go through it in a few steps. First, let's see how to evaluate $Re(e^{i\theta})$. Say we somehow obtained a state $(|0\rangle + \exp(i\theta)|1\rangle)/\sqrt{2}$. We apply H to it so that $(1+\exp(i\theta))/2$ is the term for $|0\rangle$ in the computational basis. The probability to measure 0 is therefore:

$$p(0) = \frac{1}{4}(1+e^{-i\theta})(1+e^{i\theta}) = \frac{1}{2} + \frac{1}{2}\left(\frac{e^{-i\theta}+e^{-\theta}}{2}\right) = \frac{1}{2} + \frac{1}{2}Re(e^{i\theta}) \tag{2.7}$$

Therefore, $Re(e^{i\theta}) = 2p(0) - 1$.

Now we need to get terms like those in Eqs. 2.5, generally of the form $Re(e^{i\theta}\langle 0|R^\dagger_{k,i}R_{q,j}|0\rangle)$ where $R_{k,i} = R_{N_v}...R_i\sigma_{k,i}...R_1$ and $\sigma_{k,i}$ is a Pauli matrix originating from the derivative.

First, how to get expectation values with Hermitian conjugate operators? To see how, recall that if U has eigenvalue $e^{i\phi}$, then $U^\dagger$ has eigenvalue $e^{-i\phi}$. Up to a global phase, we can get $e^{-i\phi}$ on the $|1\rangle$ state of an ancilla qubit by performing an anti-controlled U rather than controlled U.

Now let's figure out how to get $\langle 0|U^\dagger V|0\rangle = \langle U^\dagger V\rangle$. To the system register, add the ancilla qubit in a superposition state. Define $\tilde{V} = |0\rangle\langle 0|\otimes I + |1\rangle\langle 1|\otimes V$ and $\tilde{U} = |0\rangle\langle 0|\otimes U + |1\rangle\langle 1|\otimes I$. Note that these operators commute. Apply these operators to the combined ancilla and system registers and measure $\sigma^+ = \sigma_x + i\sigma_y = |0\rangle\langle 1|$ on the ancilla. That measurement will give:

$$\langle\psi_0|\tilde{V}^\dagger\tilde{U}^\dagger(\sigma^+\otimes I)\tilde{U}\tilde{V}|\psi_0\rangle = \langle\psi_0|\,|0\rangle\langle 0|\otimes U^\dagger(|0\rangle\langle 1|\otimes I)\,|1\rangle\langle 1|\otimes V\,|\psi_0\rangle \tag{2.8}$$

Taking $|\psi_0\rangle = (|0\rangle + |1\rangle)/\sqrt{2}\otimes|\psi_s\rangle$, we get

$$\langle\sigma^+\rangle = \frac{1}{2}\langle\psi_s|U^\dagger V|\psi_s\rangle \tag{2.9}$$

and so $2\langle\sigma^+\rangle = \langle U^\dagger V\rangle$. In practice, we get $\langle\sigma^+\rangle$ by measuring σ_x and σ_y and computing $\sigma_x + i\sigma_y$.

However, $\langle U^\dagger V\rangle$ was not actually what we wanted - we need the real part of it. To get that, we combine the above discussion about measuring $Re(e^{i\theta})$ with this discussion. Specifically, we apply $\tilde{U}\tilde{V}$, then $H\otimes I$, then measure the ancilla qubits. The probability to measure 0 is:

$$p(0) = \frac{1}{4}\langle\psi_s|(U^\dagger + V^\dagger)(U+V)|\psi_s\rangle \tag{2.10}$$

Recalling that $Re(U^\dagger V) = (U^\dagger V + V^\dagger U)/2$, we see that $Re(U^\dagger V) = 2p(0) - 1$.

This result is almost what we want. For the last step, first examine how to obtain $\langle T^\dagger ATB\rangle$. To compute it, let $U^\dagger = T^\dagger A$ and $V = TB$. To compute $U^\dagger$, we implement U with anti-control. Notice in the circuit with both U and V that T is both controlled and anti-controlled, meaning it is always applied! So the circuit ends up as c-B, T, c-A† with no control needed on the T operator.

Finally, we return to $\langle\psi_0|R^\dagger_{k,i}R_{q,j}|\psi_0\rangle$. Start with the ancilla in $(|0\rangle + \exp(i\theta)|1\rangle)/\sqrt{2}$. The circuit will consist of a bunch of controlled R gates for the k,i operator, then a bunch of anti-controlled gates for the q,j operator. It can be simplified as follows: assume that $k < j$. All the anti-controlled gates $R_1...\sigma_{k,i}$ can be commuted past the controlled gates to yield $R_1...R_{i-1}$ without any control, then anti-controlled $\sigma_{k,i}$ then uncontrolled $R_i...R_{j-1}$, then controlled $\sigma_{q,j}$. The rest of the gates are not needed since they act only on the system register and end up being cancelled by their Hermitian conjugates in the measurement process. This circuit, which gives the matrix elements for the matrix $A_{i,j}$ is illustrated as Fig 4 in the referenced paper and Fig 2 of Ref. [4]. The circuit for the entries of c_i is similar (given also in Fig 4 of McArdle paper) except all the R_i gates must be implemented followed by h_α at the end.

References

[1] G. Ortiz, J. E. Gubernatis, E. Knill, and R. Laflamme, "Quantum algorithms for fermionic simulations", Physical Review A **64**, 022319 (2001).

[2] R. Somma, G. Ortiz, J. E. Gubernatis, E. Knill, and R. Laflamme, "Simulating physical phenomena by quantum networks", Physical Review A **65**, 042323 (2002).

[3] A. K. Ekert, C. M. Alves, D. K. L. Oi, M. Horodecki, P. Horodecki, and L. C. Kwek, "Direct estimations of linear and nonlinear functionals of a quantum state.", Physical Review Letters **88**, 217901 (2002).

[4] Y. Li and S. C. Benjamin, "Efficient variational quantum simulator incorporating active error minimization", Physical Review X **7**, 021050 (2017).