1 Introduction

Solving eigen-pairs (eigenvalues and their corresponding eigenvectors) of Hermitian matrices is important in both theoretical and numerical aspect. For example, as for our research, we are interested in characterizing the output of the spectral initialization algorithm which calculates the leading eigenvector of a spiked matrix, under the setting that the data to generate the matrix grows proportionally in the number of data points and the dimension of each point. Theoretically the asymptotic behavior of the eigenvector sheds lights on spiked model in random matrix theory. Though we mainly focus on the theoretical characterization, the numerical experiments to validate our results are computationally hard, due to the high-dimensional structure. Practically we often use a problem-based power method or Lanczos algorithm to approximate the leading eigenvector.

Such an eigen-solving problem, in the viewpoint of quantum computation, can be equivalently paraphrased as computing the Hamiltonian ground states, which is a QMA-hard problem [1] yet there have been some quantum algorithms to achieve desired accuracy in polynomial time under special settings, e.g., [2] incorporates adiabatic evolution with phase estimation.

However, as pointed out in [3], the previous works have challenges such as requiring large quantum resources as well as high-dimensional classical optimization, which motivates to look up quantum versions of the classical eigenvalue solvers as used in our research. Fortunately, [3] provides the quantum imaginary times evolution (QITE) that is analogous to the classical power method, and quantum Lanczos (QLanczos) algorithms as well. It is not surprised that the quantum versions exponentially save space and time per iteration, and remedy the disadvantages of the previous quantum algorithms.

Our outline of the project report is as follows. First, we introduce the machinery of QITE and QLanczos in Section 2, and provide examples in special cases in Section 3. We further discuss some extensions in Section 4 and conclude our discussions at last.

2 QITE and QLanczos

Our goal is to compute nearly exactly the ground state $|\Psi_1\rangle$ of an N-qubit Hamiltonian H. Mathematically speaking, $|\Psi_1\rangle$ is the minimizer of the optimization problem

$$\min_{|\Phi\rangle} \langle\Phi| H |\Phi\rangle, \quad s.t. \ \langle\Phi|\Phi\rangle = 1.$$

If $H = \sum_{i=1}^{2N} \lambda_i |\Psi_i\rangle\langle\Psi_i|$ is the eigen-decomposition, assuming that $\lambda_1 < \lambda_2 \leq \ldots \leq \lambda_{2N} < 0$[1], then with the initial vector $|\Phi(0)\rangle = \sum_i c_i |\Psi_i\rangle$ and $c_1 \neq 0$, we have $(-H)^n |\Phi(0)\rangle = \sum_i c_i |\lambda_i|^n |\Psi_i\rangle$ which converges to $|\Psi_1\rangle$ after normalization. This is the key ingredient of power method.

---

[1]If the negativity is not satisfied, we can consider $H - m\mathbb{1}$ for large enough m. Also, such an assumption is consistent with the reality, e.g., the energy levels of the hydrogen atom are $-\frac{13.6}{n^2}$eV for positive integer n.

The continuous version is called **imaginary time evolution** and performs

$$\partial_\beta \left| \Phi(\beta) \right\rangle = -H \left| \Phi(\beta) \right\rangle \quad \Leftrightarrow \quad \left| \Phi(\beta) \right\rangle = e^{-\beta H} \left| \Phi(0) \right\rangle. \tag{1}$$

Then we have

$$\left| \Psi_1 \right\rangle = \lim_{\beta \to \infty} \frac{\left| \Phi(\beta) \right\rangle}{\langle \Phi(\beta) | \Phi(\beta) \rangle^{1/2}}, \tag{2}$$

which suggests estimating $\left| \Psi_1 \right\rangle$ by normalized $\left| \Phi(\beta) \right\rangle = e^{-\beta H} \left| \Phi(0) \right\rangle$.

2.1 Quantum imaginary time evolution

Due to the practical graphical structure, we consider a stronger locality than the local Hamiltonian discussed in class. Define a *geometric k-local Hamiltonian* $H = \sum_{m=1}^{K} h[m]$ where each $h[m]$ is a unitary acting on at most k neighbouring qubits on an underlying graph. Hereafter for simplicity we consider the graph to be a d-dimensional lattice.

To simulate $e^{-\beta H} \left| \Phi(0) \right\rangle$, write the Trotter decomposition

$$e^{-\beta H} = \left(e^{-\Delta \tau h[1]} e^{-\Delta \tau h[2]} \dots e^{-\Delta \tau h[K]} \right)^n + \mathcal{O}(\Delta \tau), \tag{3}$$

where $\Delta \tau = \frac{\beta}{n}$. Initialized with $\left| \Phi_0 \right\rangle = \left| \Phi(0) \right\rangle$, at the t-th Trotter step, we consider

$$\left| \Phi_t \right\rangle = \frac{e^{-\Delta \tau h[m]} \left| \Phi_{t-1} \right\rangle}{\sqrt{\langle \Phi_{t-1} | e^{-2\Delta \tau h[m]} | \Phi_{t-1} \rangle}}, \quad t-1 \equiv K - m \mod K. \tag{4}$$

Note that the normalization comes from the quantum view consisting of unitaries. The main idea is to regard the evolution from $\left| \Phi_{t-1} \right\rangle$ to $\left| \Phi_t \right\rangle$ as a unitary transform, solving a real Hermitian A such that

$$\left| \Phi_t \right\rangle \approx e^{-i\Delta \tau A_t} \left| \Phi_{t-1} \right\rangle \approx \left| \Phi_{t-1} \right\rangle - i\Delta \tau A_t \left| \Phi_{t-1} \right\rangle.$$

Suppose A_t acts on a domain of D qubits around the support of $h[m]$, which can be expanded by Pauli matrices on these qubits

$$A_t = \sum_{i_1 \dots i_D \in \{X,Y,Z\}} a(t)_{i_1 \dots i_D} \sigma_{i_1} \otimes \dots \otimes \sigma_{i_k} =: \sum_I a(t)_I \sigma_I. \tag{5}$$

Our goal, in specific, is to solve the optimization

$$\{a(t)_I\} = \arg \min \left(\langle \Phi_t | - \langle \Phi_{t-1} | - i\Delta \tau \langle \Phi_{t-1} | A_t^\dagger \right) \left(|\Phi_t\rangle - |\Phi_{t-1}\rangle + i\Delta \tau A_t |\Phi_{t-1}\rangle \right)$$

$$= \arg \min \left(\langle \Delta_0 | - \langle \Delta_A | \right) \left(|\Delta_0\rangle - |\Delta_A\rangle \right) \quad \left(|\Delta_0\rangle := \frac{|\Phi_t\rangle - |\Phi_{t-1}\rangle}{\Delta \tau}, |\Delta_A\rangle := -iA_t |\Phi_{t-1}\rangle \right)$$

$$= \arg \min \langle \Delta_0 | \Delta_0 \rangle - \sum_I a(t)_I \left(i \langle \Delta_0 | \sigma_I | \Phi_{t-1} \rangle - i \langle \Phi_{t-1} | \sigma_I^\dagger | \Delta_0 \rangle \right)$$

$$+ \sum_{I,J} a(t)_I a(t)_J \langle \Phi_{t-1} | \sigma_I^\dagger \sigma_J | \Phi_{t-1} \rangle. \tag{6}$$

Define

$$S_{IJ} := \langle \Phi_{t-1} | \sigma_I^\dagger \sigma_J | \Phi_{t-1} \rangle, \quad b_I := i \langle \Delta_0 | \sigma_I | \Phi_{t-1} \rangle - i \langle \Phi_{t-1} | \sigma_I^\dagger | \Delta_0 \rangle, \tag{7}$$

then the solution of the quadratic function (6) can be obtained from the linear equation

$$(\mathbf{S} + \mathbf{S}^\top)\mathbf{a}(t) = \mathbf{b}, \tag{8}$$

which can be computed by either explicitly solving the general inverse of $\mathbf{S} + \mathbf{S}^\top$ (as it may not be full-rank) or applying an iterative algorithm to get an inexact solution.

A question remains from the above least squares formulation: what should the domain size D be? There is a tradeoff between approximation accuracy and the cost of space and time. Intuitively, if the state $|\Phi_t\rangle$ is more "entangled", it will require larger support to perform a good approximation. The entanglement, or to say, the correlation is rigorously defined as follows.

Assumption 2.1 (Correlation length)**.** The correlation length of Φ_t is assumed to be upper-bounded by C for all t, which is defined as: for every $t \leq nK$ and every pair of observables A and B acting on domains separated by $\mathrm{dist}(A, B)$ sites, we have

$$C_t(A, B) := \langle\Phi_t| A \otimes B |\Phi_t\rangle - \langle\Phi_t| A |\Phi_t\rangle \langle\Phi_t| B |\Phi_t\rangle \leq \sqrt{\langle A|A\rangle\langle B|B\rangle}e^{-\mathrm{dist}(A,B)/C}. \tag{9}$$

Note that the exponential decay assumption is inherited from [4], where the correlation length C increases with β and saturates for $C \ll N$. The following theorem characterizes how well can the replacement of imaginary time evolution steps by quantum unitary updates be, as well as a suggested value of D. We postpone the proof in Appendix A, where typos and gaps in the original proof in [3] are fixed.

Theorem 2.2 (Theorem 1, [3])**.** *Under Assumption 2.1, for any $\epsilon > 0$, there are unitaries U_t each acting on $D = k(2C)^d \log^d(2\sqrt{2}nK/\epsilon)$ qubits, such that*

$$\sqrt{((\langle\Phi_{nK}| - \langle\Phi_0| U_1^\dagger \ldots U_{nK}^\dagger)(|\Phi_{nK}\rangle - U_{nK} \ldots U_1 |\Phi_0\rangle))} \leq \epsilon.$$

In a word, the unitary update can be constructed over $D = O(C^d)$ sites[2], where C is the correlation length and d is the dimension of the lattice graph. Consequently, the number of measurements and classical storage at a given Trotter step is bounded by $\exp(O(C^d))$. Note that the classical solution of the least squares problem (or solving the linear equation) has a similar exponential scaling, thus the space and time requirements are bounded by exponentials in C^d, which can be greatly reduced if there exists locality structure. We will discuss more details in section 3.

Nonetheless, it should be noted that classically the complexities are in the order of $O(\exp(n))$ and here they are only $O(poly(n))$. Such an exponential improvement shows the benefit of quantum computing.

2.2 Quantum Lanczos

Given the introduction on QITE method, we now consider the quantum version of the Lanczos algorithm, which typically converges much more quickly than imaginary time evolution in the classical setting, and practically only requires tens of iterations to achieve favorable precision.

The classical Lanczos method constructs a new set of orthogonal vectors by sequentially adding a new vector $H^t |\Phi_0\rangle$ to the subspace spanned by $\{|\Phi_0\rangle, H |\Phi_0\rangle, \ldots, H^{t-1} |\Phi_0\rangle\}$, and expresses H

---

[2]Here we omit the $\log^d(nK)$ factor in the big-O notation, which implicitly assumes that the number of trotter steps should not be too large.

in the so-called Krylov subspace. Solving the ground state in the subspace yields an estimate of the original ground state. The space and time complexities scale exponentially with the number of qubits of the Hamiltonian, which comes from storing the Lanczos vectors and the implicit dependency of the number of Lanczos iterations on the number of qubits.

The exponential bottleneck motivates us to look for the quantum version of building the Krylov space. In the following, we will show that it is efficient both in computation and in storage to consider the subspace spanned by $\{|\Phi_0\rangle, e^{-2\Delta\tau H}|\Phi_0\rangle, e^{-4\Delta\tau H}|\Phi_0\rangle, \ldots, e^{-\frac{n}{2}\cdot 2\Delta\tau H}|\Phi_0\rangle\}$. Specifically, for the n unit vectors $|\Phi_{lK}\rangle = n_l e^{-l\Delta\tau H}|\Phi_0\rangle$ where n_l is the normalization constant and $l \in [n]$, we pick all the even (or all odd) numbers of steps $\{|\Phi_0\rangle, |\Phi_{2K}\rangle, \ldots\}$ to form a basis. Observe that the picked vectors define an overlap matrix S whose elements can be computed entirely from norms:

$$S_{ll'} := \langle\Phi_{2lK}|\Phi_{2l'K}\rangle = \frac{n_{2l}n_{2l'}}{n_{l+l'}^2}\langle\Phi_{(l+l')K}|\Phi_{(l+l')K}\rangle = \frac{n_{2l}n_{2l'}}{n_{l+l'}^2}. \tag{10}$$

Thus, the overlap matrix is of size $\frac{n}{2} \times \frac{n}{2}$ after n steps of time evolution, and n_l's can be evaluated recursively as

$$n_{l+1}^2 = \frac{1}{\langle\Phi_0|(e^{-(l+1)\Delta\tau H})^\dagger e^{-(l+1)\Delta\tau H}|\Phi_0\rangle} = \frac{n_l^2}{\langle\Phi_l|e^{-2\Delta\tau H}|\Phi_l\rangle}. \tag{11}$$

Under the basis $\{|\Phi_{2lK}\rangle\}_{l=0,1,\ldots}$, the Hamiltonian satisfy

$$\overline{H}_{ll'} := \langle\Phi_{2lK}|H|\Phi_{2l'K}\rangle = \frac{n_{2l}n_{2l'}}{n_{l+l'}^2}\langle\Phi_{(l+l')K}|H|\Phi_{(l+l')K}\rangle = S_{ll'}\langle\Phi_{(l+l')K}|H|\Phi_{(l+l')K}\rangle. \tag{12}$$

Although the matrices in (10) and (12) has $\frac{n^2}{4}$ elements, there are only n unique elements to measure; more importantly, the elements are just expectations during the imaginary time evolution. Last, after obtaining the overlap matrix and the Hamiltonian matrix, one can study the eigen-equation $\overline{H}v = \lambda Sv$ and approximate the ground state by $\sum_l v_l|\Phi_{2lK}\rangle$.

The above described QLanczos shares the similar weakness of its classical counterpart: numerical unstability. If we use inexact QITE (will be discussed in Section 4) to approximate $|\Phi_{lK}\rangle$ due to a limited computational budget, then the computation of $S_{ll'}$ and $\overline{H}_{ll'}$ is no longer accurate, e.g., the overlap matrix may even be no longer positive definite! We will mention the stabilized QLanczos in Section 4 to handle this problem, as well as errors in real experiments, via ensuring non-linearity of successive vectors.

3   Special cases: simplification of QITE

Though in general the linear equation (8) costs in $\exp(O(C^d))$ scale, the complexities can be greatly reduced if imposing special structures on A_t.

1. A_t is p-local Hamiltonian. The space cost is $O(C^{dp})$ and the cost of solving the linear equation (with sparse structure) is $O(pC^dT_e)$ where T_e is the cost of computing one $a(t)_I$. This is related to the quantum algorithms of solving linear equations [5] and is out of the scope of this project.

2. A_t is geometric p-local Hamiltonian. The space cost can be further reduced to $O(pC^d)$.

3. Real Hamiltonians and states. Then $b_I = 2\Im \langle \Phi_{t-1}| \sigma_I^\dagger |\Delta_0\rangle$ defined in (7) is non-zero only if $\sigma_I^\dagger$ contains odd number of σ_Y's. By induction, one can find that the number of such operators for a domain of D qubits is $2^{D-1}(2^D - 1)$, roughly half the number of measurements needed if not assuming real Hamiltonians and states[3]. Further, as the length of **b** is half-reduced, as well as the row and column of $\mathbf{S} + \mathbf{S}^\top$, the cost of solving the linear equation can be reduced by a factor of $1/8$.

4 Variants of the algorithms

As we mentioned when introducing the two quantum methods in Section 2, with the challenge of limited resources, we will discuss the inexact QITE. Second, we will stabilize the vanilla QLanczos from noises and errors.

4.1 Inexact QITE

We have provided a suggested value of the domain size D in Theorem 2.2. However, what if we only have access to a smaller size-C neighborhood, for example, when the resources is limited? In practice, the inexact versions of the QITE and QLanczos algorithms with a smaller C can still work. Also, these algorithms heuristically outperform existing ground-state quantum algorithms which require either deep circuits or nonlinear optimization.

Further, in order to avoid numerical instabilities, we regularize $\mathbf{S}+\mathbf{S}^\top$ by adding a small diagonal regularizer $\delta\mathbb{1}$, which mitigates the effect of sampling noise.

4.2 Stabilized QLanczos

This regularization is a numerical trick without theoretical guarantee, yet works well in the tasks simulated in [3]. Instead of including all the vectors in $\{|\Phi_0\rangle, |\Phi_{2K}\rangle, \ldots\}$, we only append a new vector that is nearly orthogonal to the last one, namely, $|\langle \Phi_{2lK}|\Phi_{\text{last}}\rangle| < s$ for some pre-determined $s \in (0, 1)$, where $|\Phi_{\text{last}}\rangle$ remembers the last appended vector.

5 Conclusion remarks

In this project, we recalled two classical methods to numerically estimate the ground state of a Hermitian: imaginary time evolution (a.k.a. continuous power method) and Lanczos method, and introduced their quantum analogues. Compared to their classical counterparts, these methods exponentially reduce space and time per Trotter step or iteration, if assuming the correlation length is finite. Even when the assumption does not hold and there exists noises, the inexact and stabilized versions of the QITE and QLanczos algorithms still remain valid heuristics, which applies to the scenario when the computational budget is limited. The key point of the derivation is to transfer the non-Hermitian operation of an imaginary time step $e^{-\Delta\tau H}$ with small $\Delta\tau$ to a Hermitian operation that is natural on a quantum computer, without using ancillae and postselection.

---

[3] For each of the D qubits, one can pick from $\{\mathbb{1}, \sigma_X, \sigma_Y, \sigma_Z\}$, so there're 4^D operators in total.

References

[1] Kempe, J., Kitaev, A. and Regev, O. The complexity of the local Hamiltonian problem. *SIAM J. Comput.* 35, 1070—1097 (2006).

[2] Farhi, E., Goldstone, J., Gutmann, S. and Sipser, M. Quantum Computation by Adiabatic Evolution. Report No. MIT-CTP-2936 (2000).

[3] Motta, M., Sun, C., Tan, A.T.K. et al. Determining eigenstates and thermal states on a quantum computer using quantum imaginary time evolution. *Nat. Phys.* 16, 205–210 (2020).

[4] Hastings, M.B., Koma, T. Spectral Gap and Exponential Decay of Correlations. *Commun. Math. Phys.* 265, 781–804 (2006).

[5] Berry, D. W., Childs, A. M. and Kothari, R. Hamiltonian simulation with nearly optimal dependence on all parameters. *In 2015 IEEE 56th Annual Symposium on Foundations of Computer Science* 792–809 (2015).

[6] Brandao, F. G. and Horodecki, M. Exponential decay of correlations implies area law. *Comm. in math. phys.* 333, 761–798 (2015).

A   Proof of Theorem 2.2

The proof of Theorem 2.2 relies on the Uhlmann's theorem, which is stated as follows.

Lemma A.1 (Uhlmann's theorem). *Let $|\eta\rangle_{AB}$ and $|\nu\rangle_{AB}$ be two states with $|\cdot\rangle_A$ denoting partial traces over the complement of A. If $\| |\eta\rangle_A - |\nu\rangle_A \|_1 \leq \delta$ where $\| \cdot \|_1$ is the trace norm or the sum of the singular values, then there exists a unitary V acting on B such that*

$$\| |\eta\rangle_{AB} - (I \otimes V) |\nu\rangle_{AB} \| \leq 2\sqrt{\delta},$$

where $\| \cdot \|$ is the operation norm (ℓ_2-norm).

Our goal is to show $\| |\Phi_{nK}\rangle - U_{nK} \ldots U_1 |\Phi_0\rangle \| \leq \epsilon$, where the left-hand side can be upper-bounded by

$$
\begin{aligned}
\| |\Phi_{nK}\rangle - U_{nK} \ldots U_1 |\Phi_0\rangle \| &= \| |\Phi_{nK}\rangle - U_{nK} |\Phi_{nK-1}\rangle + U_{nK} |\Phi_{nK-1}\rangle - U_{nK} \ldots U_1 |\Phi_0\rangle \| \\
&\leq \| |\Phi_{nK}\rangle - U_{nK} |\Phi_{nK-1}\rangle \| + \| |\Phi_{nK-1}\rangle - U_{nK-1} \ldots U_1 |\Phi_0\rangle \| \\
&\leq \| |\Phi_{nK}\rangle - U_{nK} |\Phi_{nK-1}\rangle \| + \ldots + \| |\Phi_1\rangle - U_1 |\Phi_0\rangle \|.
\end{aligned}
$$

We will find a proper bound of $\| |\Phi_t\rangle - U_t |\Phi_{t-1}\rangle \|$ and its corresponding U_t from Uhlmann's theorem.

Let R_v be the region of all sites that are with at most v distance to the sites on which $h[m]$ acts (m here is the same as that in (4)). Then equipped with Assumption 2.1, Lemma 9 in [6] suggests

$$\| |\Phi_t\rangle_{R_v} - |\Phi_{t-1}\rangle_{R_v} \|_1 \leq \frac{e^{-\frac{v}{C}}}{\|e^{\Delta\tau h[m]}\|} \leq 2e^{-\frac{v}{C}}$$

for small enough $\Delta\tau \leq \frac{1}{2}$, where the last inequality uses $\|e^{\Delta\tau h[m]}\| \geq \|1 - \Delta\tau h[m]\| \geq 1 - \Delta\tau \geq \frac{1}{2}$. Thus from Lemma A.1, there exists U_t acting on R_v, such that

$$\| |\Phi_t\rangle - U_t |\Phi_{t-1}\rangle \| \leq 2\sqrt{2e^{-\frac{v}{C}}}.$$

Applying the above argument to choosing U_t's for all t, we have

$$\| \, |\Phi_{nK}\rangle - U_{nK} \ldots U_1 \, |\Phi_0\rangle \, \| \le nK \cdot 2\sqrt{2e^{-\frac{v}{C}}},$$

where the right-hand side equals ϵ when $v = 2C \log(2\sqrt{2}nK/\epsilon)$. Since the support size of the R_v-neighborhood of a k-local unitary is at most kv^d, we obtain the result for D.