Singular value decompositions in quantum computing

Matthis Lehmkuehler

17/10/2023

1. Introduction

The two big paradigms in quantum computing are changes of basis (quantum Fourier transformations being the most prominent instance of this, enabling things like phase estimation), and amplification techniques like in Grover's algorithm. All these algorithms can be unified in a rather satisfying theory as exposed in [1], [2] and [3] (see in particular this last one for the case of phase estimation and Shor's algorithm). The goal of this short note is to present the main idea behind the constructions and show how they can be used to realize the Grover speedup as a simple example.

Suppose we have n qubits and A is an $k \times l$ matrix where $k, l \leq 2^n$. Essentially, we want to implement linear algebra operations on A on a quantum computer (ideally with some speedups). Suppose that U is a unitary acting on the n qubits, $\left|\psi_1^I\right\rangle, ..., \left|\psi_l^I\right\rangle$ and $\left|\psi_1^O\right\rangle, ..., \left|\psi_k^O\right\rangle$ are two sequences of orthonormal states (the I and O stand for "in" and "out" respectively) and

$$A_{ij} = \left\langle \psi_i^O \middle| U \middle| \psi_j^I \right\rangle \quad \text{for all} \quad i \leq k, j \leq l.$$

One calls this a block embedding of A into U.

Let Π_I be the orthogonal projection onto $V_I := \left\langle \left|\psi_1^I\right\rangle, ..., \left|\psi_l^I\right\rangle \right\rangle$ and define Π_O and V_O analogously. Let us take a singular value decomposition of A, so there are orthonormal bases $\left|u_1^I\right\rangle, ..., \left|u_l^I\right\rangle$ and $\left|u_1^O\right\rangle, ..., \left|u_l^O\right\rangle$ of V_I and V_O respectively and $\lambda_1, ..., \lambda_{k \wedge l} \in [0, 1]$ such that

$$\lambda_j \delta_{ij} = \left\langle u_i^O \middle| U \middle| u_j^I \right\rangle \quad \text{for all} \quad i \leq k, j \leq l. \tag{1}$$

Our ultimate goal will be the following: Suppose that $f : [0, 1] \to \mathbb{C}$ is some function with $f(0) = 0$. The question is in which cases can we build a unitary U_f using oracle access to U and $U^\dagger$ such that

$$f(\lambda_j)\delta_{ij} = \left\langle u_i^O \middle| U_f \middle| u_j^I \right\rangle \quad \text{for all} \quad i \leq k, j \leq l.$$

i.e. we are applying the function f to the singular values of A, and how many oracle accesses are necessary to achieve this. For instance, if A is a square matrix and invertible then

$$\left(\left(A^T\right)^{-1}\right)_{ij} = \left\langle \psi_i^O \middle| U_{\lambda \mapsto 1(\lambda > 0)/\lambda} \middle| \psi_j^I \right\rangle.$$

This does not quite work since $U_{\lambda \mapsto 1(\lambda > 0)/\lambda}$ can actually not be implemented using the procedure described below but on a domain which is a positive distance away from 0 we can approximate $f_0(\lambda) = 1(\lambda > 0)/\lambda$ by a function f for which U_f can be constructed. Filling in all the details yields the singular value decomposition version of the HHL algorithm (named after Harrow, Hassidim and Lloyd who first introduced matrix inversion on quantum computers using a somewhat different algorithm), see [1] for details.

2. Construction

For $\theta \in \mathbb{R}$ we first explain how to implement the unitary $\exp(i\theta(2\Pi_I - I))$ and $\exp(i\theta(2\Pi_O - I))$. This is easy to do. Indeed, we introduce one ancilla qubit (so that we now have $n+1$ qubits in total) and let

$$F_I|\psi\rangle|i\rangle = (\Pi_I|\psi\rangle)|i\rangle + ((I - \Pi_I)|\psi\rangle)|i \oplus 1\rangle,$$

i.e. we are flipping the ancilla bit if the state $|\psi\rangle$ is in the orthogonal complement of the subspace and are leaving it unchanged if $|\psi\rangle$ is in the subspace, extending this linearly. We need to assume that we can implement this as a quantum circuit in our case. Then

$$\left(e^{i\theta(2\Pi_I - I)}|\psi\rangle\right) \otimes |0\rangle = \left(F_I\left(I \otimes e^{i\theta Z}\right)F_I\right)|\psi\rangle|0\rangle.$$

The case with Π_I replaced by Π_O is the same. By (1) we get

$$\begin{aligned} U\left|u_j^I\right\rangle - \lambda_j\left|u_j^O\right\rangle &\in V_O^\perp \quad \forall j \leq k \wedge l, \\ U\left|u_j^I\right\rangle &\in V_O^\perp \quad \forall k \wedge l < j \leq l. \end{aligned}$$

We now consider three distinct cases:

- If $k \wedge l < j \leq l$ we let $U\left|u_j^I\right\rangle =: \left|v_j^O\right\rangle \in V_O^\perp$.
- If $j \leq k \wedge l$ and $\lambda_j = 1$, then we simply have $U\left|u_j^I\right\rangle = \left|u_j^O\right\rangle \in V_O$.
- If $j \leq k \wedge l$ and $\lambda_j \neq 1$, then we define $\left|v_j^O\right\rangle \in V_O^\perp$ and then $\left|v_j^I\right\rangle$ such that

$$\begin{aligned} U\left|u_j^I\right\rangle &= \lambda_j\left|u_j^O\right\rangle + \sqrt{1-\lambda_j^2}\left|v_j^O\right\rangle, \\ U\left|v_j^I\right\rangle &= -\sqrt{1-\lambda_j^2}\left|u_j^O\right\rangle + \lambda_j\left|v_j^O\right\rangle. \end{aligned} \tag{2}$$

Inverting this equation yields

$$\begin{aligned} U^\dagger\left|u_j^O\right\rangle &= \lambda_j\left|u_j^I\right\rangle - \sqrt{1-\lambda_j^2}\left|v_j^I\right\rangle, \\ U^\dagger\left|v_j^O\right\rangle &= \sqrt{1-\lambda_j^2}\left|u_j^I\right\rangle + \lambda_j\left|v_j^I\right\rangle. \end{aligned} \tag{3}$$

Let us make some observations: First of all, the set $\left\{\left|v_j^O\right\rangle : j \leq k \wedge l, \lambda_j \neq 1\right\}$ is orthonormal. To see this, take $j, j' \leq k \wedge l$ such that $\lambda_j, \lambda_{j'} \neq 1$ and observe that

$$\begin{aligned} \delta_{jj'} = \left\langle u_j^I \middle| u_{j'}^I \right\rangle &= \left(\lambda_j\left\langle u_j^O\right| + \sqrt{1-\lambda_j^2}\left\langle v_j^O\right|\right)\left(\lambda_{j'}\left|u_{j'}^O\right\rangle + \sqrt{1-\lambda_{j'}^2}\left|v_{j'}^O\right\rangle\right) \\ &= \lambda_j\lambda_{j'}\delta_{jj'} + \sqrt{1-\lambda_j^2}\sqrt{1-\lambda_{j'}^2}\left\langle v_j^O \middle| v_{j'}^O\right\rangle \quad \text{and hence} \quad \left\langle v_j^O \middle| v_{j'}^O \right\rangle = \delta_{jj'}. \end{aligned}$$

Also we have that $\left|v_j^I\right\rangle \in V_I^\perp$ for all $j \leq k \wedge l$ with $\lambda_j \neq 1$. To prove, it is enough to show that $\left\langle v_j^I \middle| u_{j'}^I \right\rangle$ for all $j' \leq l$. This is direct if $j' > k \wedge l$, so suppose that $j' \leq k \wedge l$. If $\lambda_{j'} \neq 1$ we have

$$\begin{aligned} \left\langle v_j^I \middle| u_{j'}^I \right\rangle &= \left(\left\langle v_j^I\right|U^\dagger\right)\left(U\left|u_{j'}^I\right\rangle\right) \\ &= \left(-\sqrt{1-\lambda_j^2}\left\langle u_j^O\right| + \lambda_j\left\langle v_j^O\right|\right)\left(\lambda_{j'}\left|u_{j'}^O\right\rangle + \sqrt{1-\lambda_{j'}^2}\left|v_{j'}^O\right\rangle\right) \\ &= -\lambda_{j'}\sqrt{1-\lambda_j^2}\delta_{jj'} + \lambda_j\sqrt{1-\lambda_{j'}^2}\delta_{jj'} = 0 \end{aligned}$$

and the case $\lambda_{j'} = 1$ is similar but slightly simpler. To summarize, we have $\left|u_j^I\right\rangle \in V_I$, $\left|u_j^O\right\rangle \in V_O$ and have shown that $\left|v_j^I\right\rangle \in V_I^\perp$ and $\left|v_j^O\right\rangle \in V_O^\perp$ are normalized states and these states satisfy the equations (2) and (3).

By the definitions we also have (note that the equations involving $\left|v_j^I\right\rangle$ and $\left|v_j^O\right\rangle$ only make sense in the case where $j \leq k \wedge l$ and $\lambda_j \neq 1$ since we only defined the vectors in this case):

$$e^{i\theta(2\Pi_I - I)}\left|u_j^I\right\rangle = e^{i\theta}\left|u_j^I\right\rangle \quad \text{and} \quad e^{i\theta(2\Pi_I - I)}\left|v_j^I\right\rangle = e^{-i\theta}\left|v_j^I\right\rangle,$$
$$e^{i\theta(2\Pi_O - I)}\left|u_j^O\right\rangle = e^{i\theta}\left|u_j^O\right\rangle \quad \text{and} \quad e^{i\theta(2\Pi_O - I)}\left|v_j^O\right\rangle = e^{-i\theta}\left|v_j^O\right\rangle.$$

We see that these expressions together with (2) and (3) show what the operations U, $U^\dagger$, $R_\theta^I := \exp(i\theta(2\Pi_I - I))$ and $R_\theta^O := \exp(i\theta(2\Pi_I - I))$ do to the various states introduced above.

We can now put everything together: Take $\theta_1, \theta_1', ..., \theta_{n+1}, \theta_{n+1}' \in \mathbb{R}$ and make the following definitions

$$\begin{aligned}
U_f &:= \left(R^O_{\theta'_{n+1}} U R^I_{\theta_{n+1}}\right)\left(U^\dagger R^O_{\theta'_n} U R^I_{\theta_n}\right) \cdots \left(U^\dagger R^O_{\theta'_1} U R^I_{\theta_1}\right),\\
A_\lambda &:= \left(N_{\theta'_{n+1}} M_\lambda N_{\theta_{n+1}}\right)\left(M_\lambda^T N_{\theta'_n} M_\lambda N_{\theta_n}\right) \cdots \left(M_\lambda^T N_{\theta'_1} M_\lambda N_{\theta_1}\right) \quad \text{where}\\
N_\theta &= \begin{pmatrix} e^{i\theta} & 0 \\ 0 & e^{i\theta} \end{pmatrix}, \quad M_\lambda = \begin{pmatrix} \lambda & -\sqrt{1-\lambda^2} \\ \sqrt{1-\lambda^2} & \lambda \end{pmatrix}.
\end{aligned}$$

Then if $j \leq k \wedge l$ and $\lambda_j \neq 1$ we have $U_f\left|u_j^I\right\rangle, U_f\left|v_j^I\right\rangle \in \langle \left|u_j^O\right\rangle, \left|v_j^O\right\rangle \rangle$ and

$$\begin{pmatrix} \left\langle u_j^O\middle|U_f\middle|u_j^I\right\rangle & \left\langle u_j^O\middle|U_f\middle|v_j^I\right\rangle \\ \left\langle v_j^O\middle|U_f\middle|u_j^I\right\rangle & \left\langle v_j^O\middle|U_f\middle|v_j^I\right\rangle \end{pmatrix} = A_{\lambda_j}.$$

In particular if we let $f(\lambda) = (A_\lambda)_{11}$ then $\left\langle u_j^O\middle|U_f\middle|u_j^I\right\rangle = f(\lambda_j)$. One can check the remaining cases and verifies that indeed

$$\left\langle u_i^O\middle|U_f\middle|u_j^I\right\rangle = f(\lambda_j)\delta_{ij}.$$

So we have solved the problem if f is of the form $f(\lambda) = (A_\lambda)_{11}$ for some $\theta_1, \theta_1', ..., \theta_{n+1}, \theta_{n+1}'$. The work [1] exactly determines which functions have such a representation.

3. Grover's algorithm example

Let us consider the problem of finding $b_0 \in \{0,1\}^n$ using black-box access to the function $f : \{0,1\}^n \to \{0,1\}$ defined by $f(b) = 1(b = b_0)$. We allow oracle access to U defined by $U|b\rangle|a\rangle = |b\rangle|a \oplus f(b)\rangle$ (acting on n qubits). Note that here $U^\dagger = U$. We take $l = 1$, $k = 2^n$ and

$$|\psi_1^I\rangle = \frac{1}{\sqrt{2^n}} \sum_{b \in \{0,1\}^n} |b\rangle|0\rangle, \quad \{|\psi_1^O\rangle, ..., |\psi_k^O\rangle\} = \{|b\rangle|1\rangle : b \in \{0,1\}^n\}.$$

Then $\left\langle b, 1\middle|U\middle|\psi_1^I\right\rangle = \delta_{bb_0}/\sqrt{2^n}$ and hence $\left\langle b, 1\middle|U_f\middle|\psi_1^I\right\rangle = \delta_{bb_0} f\left(1/\sqrt{2^n}\right)$.

So if we can find an f such that U_f can be implemented with $O\left(\sqrt{2^n}\right)$ oracle queries to U and such that $|f\left(1/\sqrt{2^n}\right)| = \Omega(1)$ then we can compute $U_f|\psi_1^I\rangle$, measure in the computational basis and with probability $\Omega(1)$ the ancilla qubit will be 1 and the remaining bits give us b_0. Repeating this allows us to push the probability of never seeing the ancilla in the $|1\rangle$ state exponentially quickly (in the number of runs of the circuit) to zero, so this gives the Grover speedup as required.

Recall the definition of the Chebyshev polynomials $T_m : [-1, 1] \to [-1, 1]$ which can be implicitly be defined by

$$T_m(\cos\theta) = \cos(m\theta) \quad \text{and hence} \quad T_{2m+1}(\sin\theta) = (-1)^m \sin((2m+1)\theta).$$

In a moment, we will construct $U_{T_{2m+1}}$ using $2m+1$ oracle queries to U. But beforehand, let us explain how this yields the claim: Let m be the smallest odd integer $\geq \sqrt{2^n}$ then by the second formula $|T_m\left(1/\sqrt{2^n}\right)| = \Omega(1)$ and $m = O\left(\sqrt{2^n}\right)$ which completes the construction.

The fact that $U_{T_{2m+1}}$ can be constructed using $2m+1$ oracle queries to U follows from the discussion in the previous section and the matrix identity (noting that the top left entry on the left there is $T_{2m+1}(\cos\theta)$ as required in the setup of the previous section)

$$\begin{aligned} &\begin{pmatrix} \cos((2m+1)\theta) & -\sin((2m+1)\theta) \\ \sin((2m+1)\theta) & \cos((2m+1)\theta) \end{pmatrix} \\ &= \begin{pmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{pmatrix} \left(\begin{pmatrix} -i & 0 \\ 0 & i \end{pmatrix} \begin{pmatrix} \cos(-\theta) & -\sin(-\theta) \\ \sin(-\theta) & \cos(-\theta) \end{pmatrix} \begin{pmatrix} i & 0 \\ 0 & -i \end{pmatrix} \begin{pmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{pmatrix} \right)^m. \end{aligned}$$

This is easy to see, either by a straightforward induction or by simply observing that the product of the first three matrices in the bracket on the right implements a rotation by θ (since it is a composition of a reflection, a rotation by $-\theta$ and the inverse of the reflection) and so we are applying $2m+1$ rotations by θ in total, i.e. a $(2m+1)\theta$ rotation.

References

[1] A. Gilyén, Y. Su, G. H. Low, and N. Wiebe, "Quantum singular value transformation and beyond: exponential improvements for quantum matrix arithmetics", in *Proc. Annu. ACM Symp. Theory Comput.*, Association for Computing Machinery, 2019, pp. 193–204. doi: 10.1145/3313276.3316366. Available: https://arxiv.org/abs/1806.01838

[2] G. H. Low and I. L. Chuang, "Hamiltonian Simulation by Qubitization", *Quantum*, vol. 3, p. 163, 2019, doi: 10.22331/q-2019-07-12-163. Available: https://arxiv.org/abs/1610.06546

[3] J. M. Martyn, Z. M. Rossi, A. K. Tan, and I. L. Chuang, "Grand Unification of Quantum Algorithms", *PRX Quantum*, vol. 2, no. 4, p. 40203, 2021, doi: 10.1103/PRXQuantum.2.040203. Available: https://arxiv.org/abs/2105.02859