Variational Quantum Eigensolver

Chi-Fen Lin[a], Jia-Hua Lee[b], Huan-Hsuan Kung[b], and Yu-Lin Lu[c]

[a]Department of Chemical Engineering and Materials Science, Yuan Ze University, Taoyuan 320, Taiwan
[b]Department of Physics, National Tsing Hua University, Hsinchu 300, Taiwan
[c]Department of Physics, National Changhua University of Education, Changhua 500, Taiwan

Abstract

Variational Quantum Eigensolver, simply called VQE, is a method of hybrid quantum-classical algorithm. The main idea is to train a parametrized quantum circuit by using a classical optimizer to find the ground state energy or the low excited states of a given Hamiltonian.

1 Introduction

Variational quantum algorithm, as known as VQA, is a process of using classical optimizer to train a parametrized quantum circuit and finding the optimal parameters of the circuit.[5] It's similar to machine learning, both of them have four elements, data, model, cost function and optimizer. The difference between them is that the model of machine learning is a neural network working on classical computer, another one is a quantum circuit working on quantum computer. VQE is an example of VQA, the object is to search the ground state energy or the low excited state of a given Hamiltonian H.
In our project, we focused on finding the ground state of a Hamiltonian, that is, the lowest eigenvalue. The core concept of VQE is variational principle which can help us to obtain an upper bound of the lowest eigenvalue, mathematical interpretation will be introduced in more detail later.

2 Theory

One of the application in variational method of quantum mechanics is VQE. Before using VQE, we have to understand the mathematical background.

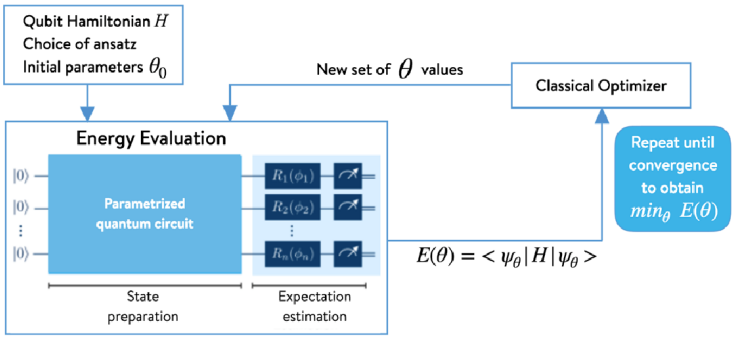


Figure 1: Simple flow chart of VQE

2.1 Mathematical Background

For a matrix A with eigenvectors $|\psi_i\rangle$ and corresponding eigenvalues λ_i, it holds the relation below which is the characteristic equation for A.[2]

$$A|\psi_i\rangle = \lambda_i|\psi_i\rangle \tag{1}$$

Then, if we have a matrix H is Hermitian, the matrix H is equal to its own conjugate transpose.

$$H = H^\dagger \tag{2}$$

Because the eigenvalue of Hermitian matrix is real, the matrix has a property $\lambda_i = \lambda_i^*$. To calculate by Quantum Machine, the matrices will be turned into a Hamiltonians. As this property, Hermitian matrices can be transformed into Hamiltonians.[2]

$$H = \sum_{i=1}^{N} \lambda_i |\psi_i\rangle\langle\psi_i| \tag{3}$$

The λ_i is the eigenvalue corresponding to the eigenvector $|\psi_i\rangle$. By matrices, we can explain the expectation value of H with

$$\langle H\rangle_\psi \equiv \langle\psi|H|\psi\rangle \tag{4}$$

By Equation (4), we subscribe the expectation value of H as[2]

$$\begin{aligned}\langle H\rangle_\psi = \langle\psi|H|\psi\rangle &= \langle\psi| \left(\sum_{i=1}^{N} \lambda_i |\psi_i\rangle\langle\psi_i|\right) |\psi_i\rangle \\ &= \sum_{i=1}^{N} \lambda_i \langle\psi|\psi_i\rangle\langle\psi_i|\psi\rangle \\ &= \sum_{i=1}^{N} \lambda_i |\langle\psi|\psi_i\rangle|^2\end{aligned} \tag{5}$$

2.2 Variational Principle

The variational principle is the core of VQE. In Quantum Mechanics, the Hamiltonian H is a matrix which describes the possible states of a physical system. Every state, expressed as eigenvector, has a corresponding eigenvalue, which represents the energy of the system. The energy value E_i for a given state $|\psi\rangle_i$ is actually the expectation value of the Hamiltonian H, like Equation(4) shown.

$$\langle\psi_i|H|\psi_i\rangle = E_i \tag{6}$$

Define the ground state energy as E_0, which we care the most, since every eigenvalue E_i is the upper bound of E_0, we can rewrite Equation (6). According to this, it is impossible to get an eigenvalue lower than the ground state energy on VQE.

$$\langle\psi_i|H|\psi_i\rangle \geq E_0 \tag{7}$$

2.3 Ansatz

As the Equation (4) and Equation (5), the second step we have to do is find a wave function, $|\psi_i\rangle$, which is called ansatz. Using Equation (5), we will get the expectation value. Therefore, finding a good ansatz is an important job. As a rule, we can normally consider that the ansatz may be cover all the condition of H we want. For example, if the expectation value of a Hamiltonian is at X-Z plane, we may use R_y Gate as the ansatz.[4] The R_y Gate means making a rotation along the y axis. That is, R_y Gate will cover all the X-Z plane. But it also gets some problems. If the expected value of Hamiltonian we get only at one side, computing all the part of R_y Gate is too heavy. The Quantum computer may not speed up the calculation.

2.4 Expectation Value

The third step, we will get the expected Expectation Value. For example, if we have a Hamiltonian is Pauli X Gate, and our mission is get the Expectation Value of Pauli X Gate. The representation of linear algebra is Equation (4). The representation of the matrix is

$$\begin{bmatrix}\alpha^* & \beta^*\end{bmatrix}\begin{bmatrix}0 & 1\\ 1 & 0\end{bmatrix}\begin{bmatrix}\alpha\\ \beta\end{bmatrix} = \begin{bmatrix}\alpha^* & \beta^*\end{bmatrix}\begin{bmatrix}\alpha\\ \beta\end{bmatrix} = \alpha^*\beta + \beta^*\alpha \tag{8}$$

With Equation (5), We know the Expectation Value show as

$$\langle X\rangle_\psi = \langle\psi|X|\psi\rangle = \sum_{i=1}^{N}\lambda_i|\langle\psi|\phi_{X\,i}\rangle|^2 \tag{9}$$

Organize the mathematical formula, we have

$$= (+1)\,|\langle\psi|\phi_{X+}\rangle|^2 + (-1)\,|\langle\psi|\phi_{X-}\rangle|^2 \tag{10}$$

Where $+1$ and -1 are eigenvalue of X Gate, ψ is our Ansatz, ϕ_X is eigenvector of X Gate. The eigenvector of X Gate are

$$\frac{1}{\sqrt{2}}\begin{bmatrix}1\\1\end{bmatrix}\ \textit{and}\ \frac{1}{\sqrt{2}}\begin{bmatrix}1\\-1\end{bmatrix}$$

In quantum computing, we can separate the eigenvalue of X Gate into $\frac{1}{\sqrt{2}}(|0\rangle + |1\rangle)$ and $\frac{1}{\sqrt{2}}(|0\rangle - |1\rangle)$. These formula can easily express by $\mathbf{H}|0\rangle$ and $\mathbf{H}|1\rangle$, which $\mathbf{H}$ is Hadamard Gate.[1] Therefore, the formula turns into

$$|\langle\psi|\mathbf{H}|0\rangle|^2 - |\langle\psi|\mathbf{H}|1\rangle|^2 = |\langle\mathbf{H}\psi|0\rangle|^2 - |\langle\mathbf{H}\psi|1\rangle|^2 \tag{11}$$

In Quantum Circuit, it is expressed as passing a wave function, apply the $\mathbf{H}$ gate, measure the Expectation Value of $|0\rangle$ and $|1\rangle$, and finally discuss the eigenvalue is 1 or -1.

3 Method

VQE involves four steps. First of all, express the given Hamiltonian in the combination of Pauli matrices.[3] Second, choose a parametrized wave function as ansatz with initial parameters θ_0 in quantum circuit. Third, measure the expectation value $E(\theta) = \langle\psi_\theta|H|\psi_\theta\rangle$ on quantum computer. The last, use classical optimizer to obtain the minimum $E(\theta)$.
In our project, we worked on a given Hamiltonian H in Equation (12). We will demonstrate the steps of VQE by taking H as an example.

$$H = \begin{pmatrix}1 & 0 & 0 & 0\\ 0 & 0 & -1 & 0\\ 0 & -1 & 0 & 0\\ 0 & 0 & 0 & 1\end{pmatrix} \tag{12}$$

3.1 Transform the Hamiltonian

For any given Hamiltonian, one can transform into a combination of Pauli matrices.[3] Since we once knew the method of measuring Pauli matrices, the problems were smoothly solved.
Look at Equation (12), it was not difficult to find that the matrix only has values on diagonal lines. Luckily, both $\sigma_x \otimes \sigma_x$ and $\sigma_y \otimes \sigma_y$ are anti-diagonal matrices, and $\sigma_z \otimes \sigma_z$ is a diagonal matrix. Therefore, we can transform H into the forms of Pauli matrices combinations.

$$H = -\frac{1}{2}\sigma_x \otimes \sigma_x - \frac{1}{2}\sigma_y \otimes \sigma_y + \frac{1}{2}\sigma_z \otimes \sigma_z + \frac{1}{2}I_{4\times 4} \tag{13}$$

3.2 Determine the Ansatz

As mentioned above, we have to design a proper wave function which can cover the lowest state and improve the efficiency at the same time. To make sure the wave function can find the ground state energy of a given Hamiltonian, there is a general form of ansatz for the states of two qubits.

$$|\psi\rangle = (U' \otimes U'')(CX_{0,1})(U \otimes I_{2\times 2})|0\rangle|0\rangle \tag{14}$$

Where U is the general form of a single qubit unitary and $CX_{0,1}$ is CNOT gate.[7]
Equation (14) obtained by Schmidt decomposition theorem.[8] Since any bipartite quantum state $|\psi\rangle$ can be expressed as a general 2-qubit state.

$$|\psi\rangle = \sum_{\alpha} \sigma_\alpha(|1\rangle \otimes |0\rangle) \tag{15}$$

By the Schmidt decomposition and the singular value decomposition theorem,[8][9] we can change Equation (15) into the form of Equation (26) with $|\alpha_i\rangle = |\beta_i\rangle = |0\rangle$. Since one can find a unitary matrix U to rotate the single-state $|0\rangle$ to $|1\rangle$,[7] we can rewrite Equation(15).

$$\begin{aligned} |\psi\rangle &= \sum_{\alpha} \sigma_\alpha[(\sum_{a} U_{a\alpha}|0\rangle) \otimes ({I_{2\times 2}}^T|0\rangle)] \\ &= \sum_{\alpha,a} \sigma_\alpha U_{a\alpha} I_{2\times 2}|0\rangle|0\rangle \\ &= \sum_{i,j} \lambda_{i,j}|0\rangle|0\rangle \end{aligned} \tag{16}$$

According to the first line of Equation (16), one can tell that any 2-qubit state can be formed from two qubits with state $|0\rangle$ by applying a U gate on one of the qubits, a CNOT gate to create entanglement, and U gates on both qubits.
Back to our given Hamiltonian in Equation (12), we can define the general form as the ansatz, however, there are nine parameters in total, which is inefficient for VQE since the unreasonable number of parameters will slow down the speed of optimizing in classical computer, therefore, design an ansatz as simple as possible. If we have adequate knowledge in the given Hamiltonian, we can reduce the numbers of parameters effectively. In our case, we designed the ansatz as

$$|\psi\rangle = (R_x(\theta) \otimes I_{2\times 2})(CX_{0,1})(\mathbf{H} \otimes I_{2\times 2})|0\rangle|0\rangle. \tag{17}$$

3.3 Measure the Expectation Value

In quantum circuit, after the qubits pass the ansatz, we need to apply different gate to measure the expectation. How to choose the gate is based on the expectation of what we want to ask for. In our project, as the Equation (13), σ_x is X gate, σ_y is Y gate, σ_z is Z gate in quantom circuit. So, we need to measure the expectation value of $X \otimes X$, $Y \otimes Y$, $Z \otimes Z$, and I. As the Equation(9), (10), and (11), we have known that

$$\langle X\rangle_\psi = |\langle \mathbf{H}\psi|0\rangle|^2 - |\langle \mathbf{H}\psi|1\rangle|^2 = P_{\psi_x}(0) - P_{\psi_x}(1) \tag{18}$$

Where $\psi_x = \mathbf{H}\psi$, $P_{\psi_x}(0)$ is the probability that after passing the wave function, the measurement of qubit is 0, and $P_{\psi_x}(1)$ is the probability that the measurement of qubit is 1. We can use the same mathematic method to get

$$\langle Y\rangle_\psi = |\langle \mathbf{H}S^\dagger\psi|0\rangle|^2 - |\langle \mathbf{H}S^\dagger\psi|1\rangle|^2 = P_{\psi_x}(0) - P_{\psi_x}(1) \tag{19}$$

Where $\psi_y = \mathbf{H}S^\dagger\psi$. [11]
Notice we should apply Hadamard Gate to the qubit before measurement the expectation of Pauli X. By mathematic theorem[10] , we can easily imply that if we want to measurement the expectation of $(X \otimes X)$, we should apply Hadamard Gate on both two qubits. That is

$$\begin{aligned} \langle X \otimes X\rangle_\psi &= |\langle\psi'|00\rangle|^2 + |\langle\psi'|11\rangle|^2 - |\langle\psi'|01\rangle|^2 - |\langle\psi'|10\rangle|^2 \\ &= P_{\psi'}(00) + P_{\psi'}(11) - P_{\psi'}(01) - P_{\psi'}(10) \end{aligned} \tag{20}$$

Where I've define $\psi' = (\mathbf{H} \otimes \mathbf{H})\psi$. We can also get

$$\begin{aligned} \langle Y \otimes Y \rangle_\psi &= |\langle \psi'' | 00 \rangle|^2 + |\langle \psi'' | 11 \rangle|^2 - |\langle \psi'' | 01 \rangle|^2 - |\langle \psi'' | 10 \rangle|^2 \\ &= P_{\psi''}(00) + P_{\psi''}(11) - P_{\psi''}(01) - P_{\psi''}(10) \end{aligned} \tag{21}$$

Where I've define $\psi'' = (\mathbf{H}S^\dagger \otimes \mathbf{H}S^\dagger)\psi$. And

$$\begin{aligned} \langle Z \otimes Z \rangle_\psi &= |\langle \psi | 00 \rangle|^2 + |\langle \psi | 11 \rangle|^2 - |\langle \psi | 01 \rangle|^2 - |\langle \psi | 10 \rangle|^2 \\ &= P_\psi(00) + P_\psi(11) - P_\psi(01) - P_\psi(10) \end{aligned} \tag{22}$$

And it is clearly to get the expectation of I is 1 without measurement since it just sum all of the probability .

$$\begin{aligned} \langle I \rangle_\psi &= |\langle \psi | 00 \rangle|^2 + |\langle \psi | 11 \rangle|^2 + |\langle \psi | 01 \rangle|^2 + |\langle \psi | 10 \rangle|^2 \\ &= P_\psi(00) + P_\psi(11) + P_\psi(01) + P_\psi(10) \\ &= 1 \end{aligned} \tag{23}$$

In our quantum circuit, it involes ansatz and the gate for expectation measurement. There are some steps to measure the probability. First, we design the quantum circuit and measure the qubits, we can obtain the value of qubits. Second, we run this circuit for serveral times, we called it shot number, and get the different result of measurement. The last, we count the different result by the computer and finally can get the probability we want. Back to our given Hamiltonian in Equation (12), we want to obtain the expectation of H. We can rewrite Equation(13)

$$\langle H \rangle_\psi = -\frac{1}{2}\langle X \otimes \ X \rangle_\psi - \frac{1}{2}\langle Y \otimes Y \rangle_\psi + \frac{1}{2}\langle Z \otimes Z \rangle_\psi + \frac{1}{2}\langle I_{4\times 4} \rangle_\psi \tag{24}$$

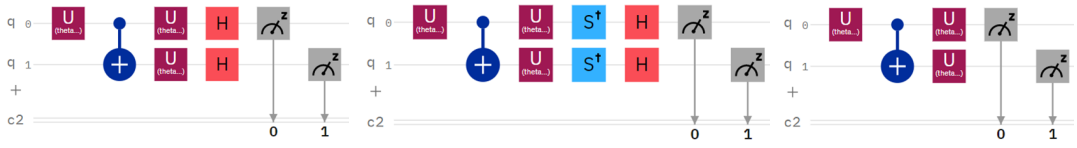
As mention above the method, we can obtain the expectation.

3.4 Optimization

Notice the Equation(14), the wave function is depand on nine parameters. Because of it, we will get different expectation with different value of parameters. In our project, our purpose is to get the lowest expectation, so we need the classical optimizer to help us find it.

4 Results

Our target Hermitian is in Equation(12). Since the eignvalue of given Hermitian are 1 and -1, the ground state energy E_0 is -1. Due to Equation(7), the theoretical value of lowest expectation is -1. Accroding the VQE method we mention above, Figure 2 shows the quantum circuit we design.



(a) $\langle X \otimes \ X \rangle_\psi$ (b) $\langle Y \otimes \ Y \rangle_\psi$ (c) $\langle Z \otimes \ Z \rangle_\psi$

Figure 2: Quantum Circuit

As the Equation(24) and the general form of ansatz is in Equation(14), there are nine parameters. We use the optimizer to help us find the lowest expectation of Hermitian in Equation(12). We measure three times with shot number = 10000. The result are respectively -0.99980, -0.99900, and -0.99940. The average is -0.99940. The error is 0.060%. The average execution time is 12.1 s.
Then, we notice that both of the shot number and the number of parameters in ansatz will influence the error and the execution time. To prove this argument, we design a ansatz in Equation(17) which is only has one parameter and change the shot number to compare it. Each one is averaged after we measure three times. The result are shown in Tabel 1.

Number of parameters	shot number	expectation	error	execution time(s)
9	10000	-0.99940	0.060 %	12.1
	100000	-0.99994	0.006 %	66.0
1	10000	-1	0.000 %	1.75
	100000	-1	0.000 %	10.97

Table 1: The impact of different number of parameters and shot number on the expectation

Note that ansatz in Equation(17) which is only has one parameter is just we guass since we have adequate knowledge in the given Hamiltonian. This is not the formal process. We take this ansatz is just want to show how important designing the ansatz with lower number of parameters is. As we can see, at the same shot number, lower number can reduce the error and the execution time obviously. Beside, we can also see that the shot number will influence the error and execution time. We'll discuss later.

5 Discussion

We found in the implementation that the number of shots will affect the error of the lowest eigenvalue. Here we use another matrix as an example.

$$M = \begin{pmatrix} -0.2524859 & 0.18121 \\ 0.18121 & -1.8318639 \end{pmatrix} \tag{25}$$

We set the number of shots from 10,000 to 1 million, and measure the lowest eigenvalue three times repeatedly, then take the average, and finally calculate the error for comparison.

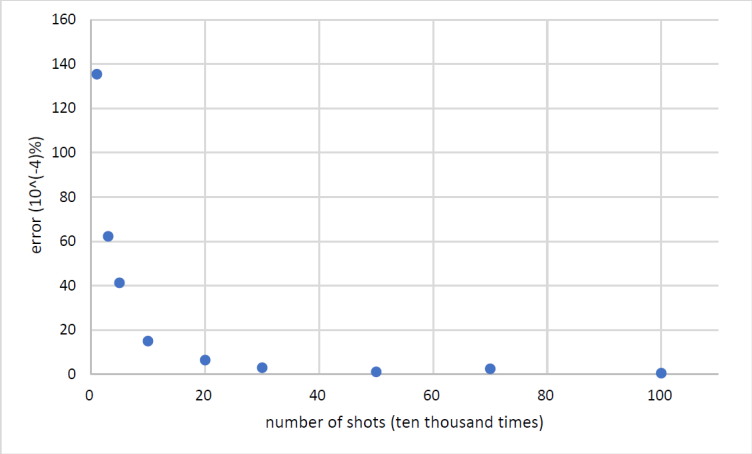


Figure 3: The relationship between the number of shots and the error

According to our observations and experiments, the accuracy of this method is related to the number of shots. It can be seen from the figure that the larger the number of shots, the smaller the error of the measured data.

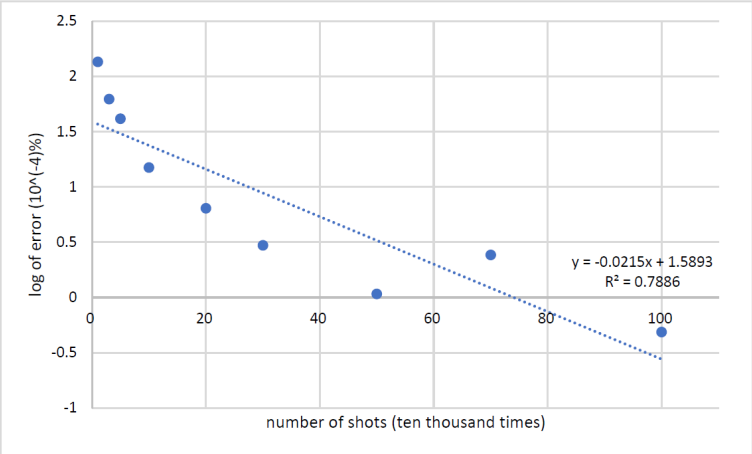


Figure 4: The relationship between the number of shots and the log of error

Then we take the log of error and get the regression line. According to the references section 2.3 [6], the number of shots required for different accuracy can be calculated through formulas. When the accuracy is high, the number of shots should be correspondingly larger.

6 Acknowledgments

We are very grateful to Teacher, Cheng-Lin Hong, for teaching us all of the knowledgement, and also for his helpful comments on this paper. We also thank all of the departments and teachers who helped organize this activity, Quantum Technology Summer School, so we can get knowledgement and finish this paper.

References

[1] $\mathbf{H} = \frac{1}{\sqrt{2}}\begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix}$

[2] Qiskit Textbook – Simulating Molecules using VQE
`https://qiskit.org/textbook/ch-applications/vqe-molecules.html`

[3] $\sigma_x = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \quad \sigma_y = \begin{pmatrix} 0 & i \\ -i & 0 \end{pmatrix}, \quad \sigma_z = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}$

[4] $R_y(\theta) = \begin{pmatrix} \cos\frac{\theta}{2} & -\sin\frac{\theta}{2} \\ \sin\frac{\theta}{2} & \cos\frac{\theta}{2} \end{pmatrix}$

[5] Musty Thoughts – Variational Quantum Algorithms-How do they work?
`https://www.mustythoughts.com/vqas-how-do-they-work`

[6] Strategies for quantum computing molecular energies using the unitary coupled cluster ansatz
`https://iopscience.iop.org/article/10.1088/2058-9565/aad3e4/pdf`

[7] Quantum logic gate:

1. U gate: for any rotation in space, one can describe by three angles called Euler angle in classical mechanics
$U = u_3(\theta, \phi, \lambda) = \begin{pmatrix} cos(\frac{\theta}{2}) & -e^{i\lambda}sin(\frac{\theta}{2}) \\ e^{i\phi}sin(\frac{\theta}{2}) & e^{i(\lambda+\phi)}cos(\frac{\theta}{2}) \end{pmatrix}$

2. CNOT or CX gate: a 2-input quantum gate, the first index is controlled qubit and the second one is the target

$$CX_{0,1} = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 \end{pmatrix}$$

[8] Schmidt decomposition theorem:
For bipartite quantum states, let $|\psi\rangle_{AB}$ be any quantum state of a composition system comprising an m dimensional system A with orthonormal bases $\{|\alpha_1\rangle, ..., |\alpha_m\rangle\}$ and n dimensional system B with bases $\{|\beta_1\rangle, ..., |\beta_n\rangle\}$. Let $d = min\{m, n\}$ and there are non-negative real numbers $\lambda_1, ..., \lambda_d$ called the Schmidt coefficients of $|\psi\rangle$, such that

$$|\psi\rangle = \sum_{i=1}^{d} \lambda_i |\alpha_i\rangle |\beta_i\rangle \tag{26}$$

[9] Singular value decomposition:
Let A is a $m \times n$ matrix, U and V are $m \times m$ and $n \times n$ unitary matrices respectively, that is, $U^T = U^{-1}$, $V^T = V^{-1}$ and $\sum$ is a $m \times n$ rectangular diagonal matrix with non-negative real numbers. Then, A can be expressed in the form $U \sum V^T$.

[10] Texxtbook - Matrix Analysis for Scientists and Engineers, Alan J. Laub. Chapter 13, Theorem 13.12, https://archive.siam.org/books/textbooks/OT91sample.pdf
Let a $n \times n$ matrix A have eigenvalues λ_i, and let a $m \times m$ matrix B have eigenvalues μ_j. Then the eigenvalues of $A \otimes B$ are $\lambda_1\mu_1$,... $\lambda_1\mu_m$, $\lambda_2\mu_1$,...,$\lambda_2\mu_m$,...,$\lambda_n\mu_m$. x_i are linearly independent right eigenvectors of A corresponding to λ_i. z_j are linearly independent right eigenvectors of B corresponding to μ_j. Then $x_i \otimes z_j$ are linearly independent right eigenvectors of $A \otimes B$ corresponding to $\lambda_i\mu_j$.

[11] $S^\dagger = \begin{pmatrix} 1 & 0 \\ 0 & -i \end{pmatrix}$