Schmidt Decomposition in Quantum Information Theory

Using the Singular Value Decomposition to Initialize Quantum States

April Reisenfeld and Charlie Bushman

1 Quantum Information Theory

1.1 Information Theory

After the development of the telegraph and telephone in the late 1800s, the ability to characterize and quantify communication signals became relevant to the workings of everyday life. Systematic methods through which signal analysis could be conducted, however, did not become formalized until the 1920s when Harry Nyquist and others published mathematical techniques for measuring data transmission rates [1]. The next leap in information theory, and the birth of "modern" information theory, was initiated by electrical engineer Claude Shannon who formalized information as a characteristic of messages that is completely separate from the meaning of messages. In other words, he defined information theory as applicable to questions of communication transmission only, whereby questions regarding semantics cannot be addressed [1]. The kinds of technical problems associated with the concepts, parameters, and rules governing communication transmission are robust, resulting in vast applicability of the theory to many disciplines [2].

The success of Shannon's classical information theory in analyzing various problems was dependent on the simplicity of his communication model. In this model, a message being transmitted is first encoded into a signal, then sent through a noisy channel, decoded back into a message, and finally received (See Fig. 1).

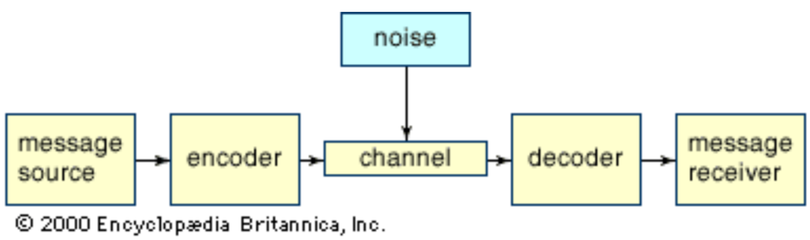

Figure 1: Figure of Shannon's communication model. The arrows indicate signal transmission direction [1].

Once simplified into a discrete model, Shannon was able to generalize and reduce the huge, complex problem of communication into essentially two smaller problems: message encryption and transmission. Shannon's model of communication provides insight into not only the transmission of signals, but also the nature of information itself. Once separated from semantics, information can be quantized, measured, and analyzed formally through mathematics. In the encoding process, information is transformed into a generalized signal in the form of binary digits. These binary digits, known as "bits," are units of information that represent one of two possible distinct states. Due to the wide variety of physical systems that exhibit two possible distinct states, for example a coin flip, information encoding and storage can occur within many contexts. Shannon realized that the amount of information within an encoded message, once converted into bits, is dependent on the amount of unpredictability of the bit string; the more unpredictable is it the more information it has. Messages that have repeating patterns can be condensed into shorter bit strings, thereby possessing less information than ones that are completely random. For example, the bit string

$$0000011111$$

contains less information than the string

$$0100010010$$

since the former is more predictable and able to be compressed into a statement such as $5(0)+5(1)$. Due to the similarity of its thermodynamic counterpart, Shannon decided to name the quantity that characterizes this randomness as entropy, a scalar quantity used to measure the amount of choice involved in the selection of an outcome, or the degree of uncertainty in the outcome.

Shannon's paradigm shift of transforming Boolean logic that specifies the state of a system (True, False) into binary digits revolutionized computing, as it allowed data to be encoded, compressed, and sent across noisy channels. Information contracted a new meaning in which it is connected to objective probabilities, allowing for its measurement and quantification. However, the macroscopic world characterized by classical mechanics is not the only scale which possesses binary state systems. The quantum world of electron spin and photon polarization can also be modeled using Shannon's information theory as "qubits," or quantum bits, a field of work that has developed into quantum information theory. In the late 1900s, quantum computing began building upon the theories of quantum information to create methods for improving computation speeds to tackle problems that were too complex for classical computers [4]. The race to build the first quantum computer has held companies such as Google and IBM in competition for decades due to its promise of solving presently-unsolvable problems. The superiority of quantum computing utilizes quantum information theory to build quantum circuits that exploit the foundational components of quantum mechanics.

1.2 Quantum Mechanics

Information theory emphasizes and exploits the close connection between information and probability, whereby the values we assign to probabilities depend on available information and information is a function of probabilities [5]. Due to the probabilistic nature of quantum mechanics in which experimental results can be only probabilistically predicted both in practice and in theory, information theory becomes an inevitable tool for understanding and utilizing its physical laws. The quantum mechanical rules for dynamics and measurement along with the principles of superposition and entanglement provide quantum information theory with its distinct features [5]. Tasks that are either impractical or impossible in the classical domain become possible due to these rules. The foundational mathematical concepts in quantum mechanics necessary to formulate quantum information theory and a notion of qubits include quantum states and superposition, measurement and observables, and entanglement.

The mathematical formulation of quantum mechanics is based around the idea of topologically separable Hilbert spaces in which the state of a physical system can be completely specified by its state vector. This state vector is a column vector in the Hilbert space known as a "ket", denoted $|\psi\rangle$. Its transpose is called a "bra" and is denoted $\langle\psi|$. This notation comes from the combination of "bra" and "ket" as "braket", $\langle\psi|\psi\rangle$, which represents an inner product of the vector $|\psi\rangle$ with itself. Similarly, the outer product can be represented as $|\psi\rangle\langle\psi|$. As in Euclidean vector space, Hilbert space quantum state vectors obey linearity and can be superposed to create new states. If $|\psi_1\rangle$ and $|\psi_2\rangle$ are states, then

$$|\psi\rangle = a_1|\psi_1\rangle + a_2|\psi_2\rangle \tag{1}$$

is also a state of the system, where a_1, a_2 are complex numbers [5]. The principle of superposition is one of the fundamental differences between quantum and classical mechanics; systems can exist in indefinite states before measurement. The scalar quantities a_1, a_2 are determined by finding the overlap of each state $|\psi_1\rangle, |\psi_2\rangle$ with the superposed state $|\psi\rangle$ as $\langle\psi_1|\psi\rangle, \langle\psi_2|\psi\rangle$. Under conditions of normalization consistent with the inner product of two vectors, we find the ability to probabilistically predict the state of a system with these coefficients known as probability amplitudes. The square of these amplitudes gives the probability of observing one of the two superposed states when the system is measured. In general, we can define a state as

$$|\psi\rangle = \sum_n a_n|\psi_n\rangle, \tag{2}$$

where, if $|\psi\rangle$ is normalized and $|\psi_n\rangle$ are orthonormal, then

$$\sum_n |a_n|^2 = 1, \tag{3}$$

which is consistent with the probability amplitude interpretation of coefficients a_n [5].

A measurement of a quantum system is performed through the action of an operator on a quantum state. Specifically, an observable quantity is produced by the action of a Hermitian operator $\hat{A}^\dagger = \hat{A}$ such that

$$\hat{A}|\lambda_n\rangle = \lambda_n|\lambda_n\rangle, \tag{4}$$

where $|\lambda_n\rangle$ are eigenvectors and the eigenvalues λ_n are observables. The mean value of the action of A from measurements on an ensemble of identical systems is the expectation value

$$\bar{A} = \langle\hat{A}\rangle = \sum_n \lambda_n|a_n|^2 = \langle\psi|\hat{A}|\psi\rangle. \tag{5}$$

Other kinds of operators beyond Hermitian may also act on quantum systems, such as unitary operators $\hat{U}\hat{U}^\dagger = I = \hat{U}^\dagger\hat{U}$. Whereas Hermitian operators are information-extraction elements in quantum computing, unitary operators are information-processing elements. Information processors can be thought in terms of a design of suitable unitary operators and implementations for the appropriate quantum evolution [5]. The unitary evolution of a system is analogous to a rotation in a space spanned by the state vectors (See Fig. 2). Consequently, a quantum system can be initialized to a desired state using a sequence of unitary operators.

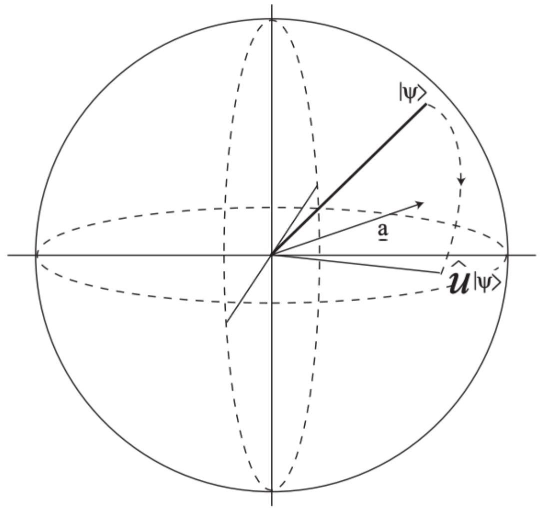

Figure 2: Schematic of the action of a unitary operator $\hat{U}$ on a quantum state $|\psi\rangle$. The state vector is rotated in the space spanned by the state vectors $|\psi_n\rangle$ from Eq. 2 about an axis a. This state representation is known as a Bloch sphere.

1.3 Entanglement and Schmidt Decomposition

Quantum information theory utilizes these quantum mechanical principles to create and manipulate the quantum information units known as qubits. The two-state systems initially motivated by Shannon for classical computing can be created using a quantum system that possesses two quantum states, $|0\rangle$ and $|1\rangle$. For example, the polarization of a photon, the spin orientation of a spin-half particle, or a pair of electron energy levels are two-state quantum systems that can represent a qubit [5]. Not only can these qubits be initialized to a single state as in classical computing, but as we have previously seen, they can be initialized to any superposition of the two states, $a_0\,|0\rangle + a_1\,|1\rangle$. These qubits can be combined and acted on by operators for the purpose of information processing and extraction. Specifically, a quantum system composed of multiple qubits, each of which is prepared in state $|0\rangle$, is written as

$$|\psi\rangle = |0\rangle \otimes |0\rangle \otimes ... \otimes |0\rangle\,, \tag{6}$$

where $\otimes$ is a tensor product, and the number of $|0\rangle$ terms is the number of qubits in the system [5]. This multiple qubit states can also be written more succinctly as $|00...0\rangle$. Once prepared, we can act on these systems in many ways. Operators can be used to transform single qubits, or transformations can be used to couple multiple qubits to create entangled states.

Entanglement is the final component of quantum mechanics that presents quantum systems with unique properties for computing. It specifies correlations between two or more quantum systems. For a state specifying a composite quantum system composed of two individual states $|\lambda\rangle$, $|\phi\rangle$, from systems a, b we can specify

$$|\psi\rangle = |\lambda\rangle_a\,|\phi\rangle_b\,, \tag{7}$$

which is simply a direct product of the individual states. We know from Eq. 1 that any superposition of product states is also an allowed state for our composite system. If we have a composite state that is a superposition of product states,

$$|\psi\rangle = \frac{1}{\sqrt{2}}(|0\rangle_a |0\rangle_b + |1\rangle_a |1\rangle_b), \tag{8}$$

such that it is not factorable into states for the two systems a, b, then we say this state is entangled. The consequence of this property is that since the two systems a, b, which may represent two qubits, cannot be separated, information about one tells us information about the other. For example, if we know that qubit a of the composite system shown in Eq. 8 is in state $|0\rangle$, then we know that qubit b must also be in state $|0\rangle$. The ability to entangle qubits allows for increased speed of computing since more information can be produced with less computation. Due to the relevance of entangled states for the creation of quantum circuits composed of many qubits, it is extremely important to be able to quantify the amount of entanglement present in any given composite state. The form of quantum states that allows for this quantization is known as the Schmidt decomposition, named after mathematician Erhard Schmidt, which is essentially a restatement of the familiar singular value decomposition.

2 Mathematical Formulation

The Schmidt decomposition represents a quantum state vector as a tensor product of two quantum vector spaces, such as in Eq. 7. If we consider the bipartite quantum state $|\psi\rangle$, that is given as

$$|\psi\rangle = \sum_{a,b} \psi_{ab} |a, b\rangle\,, \tag{9}$$

where $|a, b\rangle$ is the standard basis of the composite system $\mathcal{H}_A \otimes \mathcal{H}_B$, and ψ_{ab} are coefficients. Since ψ_{ab} can also be represented as a matrix of coefficients, we can take the singular value decomposition (SVD) of this matrix. The SVD gives us the form of each coefficient ψ_{ab} as,

$$\psi_{a,b} = \sum_{\alpha} \sigma_\alpha U_{a\alpha} V^*_{b\alpha}. \tag{10}$$

Since the state $|\psi\rangle$ must be normalized according to Eq. 3, we know that

$$\begin{aligned} \sum_{ab} |\psi_{ab}|^2 &= 1, \\ \sum_{\alpha=1}^{r} \sigma_\alpha^2 &= 1. \end{aligned} \tag{11}$$

With substitution, we can rewrite Eq. 9 as

$$\begin{aligned} |\psi\rangle &= \sum_{a,b,\alpha} \sigma_\alpha U_{a\alpha} V^*_{b\alpha} |a, b\rangle \\ &= \sum_{\alpha} \sigma_\alpha \big[\sum_a U_{a\alpha} |a\rangle \big] \otimes \big[\sum_b V^*_{b\alpha} |b\rangle \big]. \end{aligned} \tag{12}$$

If we define

$$
\begin{aligned}
|\alpha_a\rangle &= \sum_a U_{a\alpha} |a\rangle , \\
|\alpha_b\rangle &= \sum_b V^*_{b\alpha} |b\rangle ,
\end{aligned}
\tag{13}
$$

then we obtain the Schmidt decomposition form of our state vector:

$$
|\psi\rangle = \sum_\alpha \sigma_\alpha |\alpha_a\rangle \otimes |\alpha_b\rangle , \tag{14}
$$

where the Schmidt coefficients are the squares of the singular values, $\lambda_\alpha = \sigma_\alpha^2$ [6]. In this form, $|\alpha_a\rangle , |\alpha_b\rangle$ are orthonormal bases for the two systems a, b. Each state $|\alpha_a\rangle$ of qubit a is uniquely associated with a state $|\alpha_b\rangle$ of qubit b [5].

This form of the quantum state $|\psi\rangle$ allows for easier computation since we only have a single sum, rather than the double sum in Eq. 9. Additionally, we have introduced Schmidt coefficients λ_α which provide information about the entanglement of the system. Since we know from Eq. 8 that an entangled state is one that cannot be factored, we can see from this form that a state with more than one singular value will be a product state [6]. The quantification of the number of Schmidt coefficients is called the Schmidt rank, where entangled states have a rank greater than one. The Schmidt coefficients also contain information about the degree of entanglement, and it can be shown that the state with maximum entanglement is one in which all Schmidt coefficients are equal [6]. Quantifying the degree of entanglement, however, is not the only utility of the Schmidt decomposition. Representing the quantum state $|\psi\rangle$ as a product of unitary matrices U, V and singular values allows us to recreate $|\psi\rangle$ with a series of unitary matrices operating on a two-qubit system.

3 Qubit State Initialization

3.1 Developing the Theory

The initialization of qubit states on NISQ (noisy intermediate-scale quantum) computers, a category which all current quantum hardware falls into, is a nontrivial problem in quantum computation. Unlike in classical computing where we can begin any computation by initializing variables however we like, typical quantum hardware will initialize an n-qubit state to $|0\rangle$. It is then up to the programmer to find the correct unitary matrix $T \in U(2^n)$ that will rotate $|0\rangle$ to the desired initial position on the Bloch sphere(s), $|\psi\rangle$, as in Fig. 2.

There are many operators, known as quantum logic gates, that can be used to achieve this state initialization. One ubiquitous gate that is an essential component of gate-based quantum computing is the Controlled-NOT (CNOT) gate. The CNOT gate operates using two qubits, with one as control and the other as target, flipping the state of the target qubit if and only if the control qubit state is $|1\rangle$. Using these gates along with those from the Schmidt decomposition of our coefficient matrix defined in the previous section, we can create a combined operator T that performs the desired initialization operations:

$$T = (U \otimes V)(C^1_{n+1} \otimes C^2_{n+2} \otimes ... \otimes C^n_{2n})(B \otimes I), \tag{15}$$

where U and V are the left and right singular matrices from Eq. 10, C^j_i is a CNOT gate with control j and target i, and B is a unitary matrix with the singular values σ_i as entries of the first column [7]. The implementation of this combined operator T in a quantum circuit is seen in Fig. 3. caveat of this definition is that the T defined above is an operator in $U(2^{2n})$ meaning that it initializes a $2n$-qubit state. This is often inconsequential, but in the cases where an odd number of qubits is needed there are other methods for doing that. The form of T given in Eq. 15 is derived from the Schmidt decomposition of the desired state $|\psi\rangle$ given in Eq. 12. However, with a desired state $|\psi\rangle$ in mind, the operators U, V, and B need to be defined in order to operate with T. U, V we know can be found by applying the SVD to the coefficient matrix ψ_{ab}, however finding a unitary matrix B is slightly more involved.

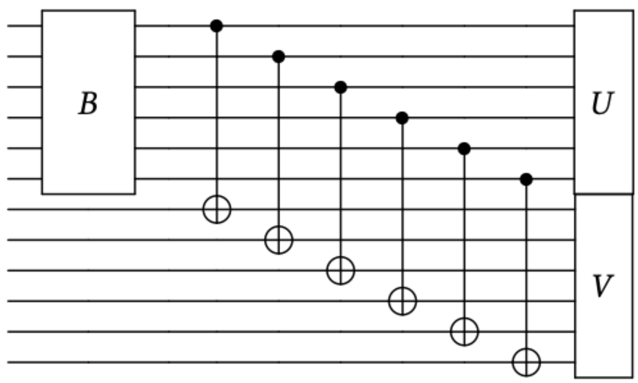

Figure 3: T operator acting on a six-qubit system for initialization to a desired state. U, V are unitary operators derived from the SVD of the coefficient matrix, and B is a unitary operator derived below. In between the unitary operators are six CNOT gates, where the line connects a target qubit (+) with a control qubit (solid).

Let's start from a seemingly arbitrary place where we assume B takes the form given by Eq. 16 and see if we can force it to satisfy our requirements of being unitary and having a first column given by $|\Sigma\rangle = \sum_n \sigma_n |n\rangle$.

$$\begin{aligned} B &= 2|q\rangle\langle q| - I, \\ ||q|| &= 1. \end{aligned} \tag{16}$$

First we can prove that this is unitary by showing that $B^\dagger B = BB^\dagger = I$:

$$\begin{aligned} B^\dagger B &= (2|q\rangle\langle q| - I)^\dagger(2|q\rangle\langle q| - I) \\ &= 4|q\rangle\langle q|\,|q\rangle\langle q| - 4|q\rangle\langle q| + I \\ &= I, \end{aligned} \tag{17}$$

and the same logic holds for $BB^\dagger$. So we know that B is unitary. Next, $|q\rangle$ must be solved for such that the first column of B is $|\Sigma\rangle$, which just means solving for $|q\rangle$ in

$$(2\,|q\rangle\,\langle q| - I)\,|1\rangle = |\Sigma\rangle\,. \tag{18}$$

With some relatively simple manipulations, this comes out to,

$$|q\rangle = \frac{|1\rangle + |\Sigma\rangle}{\sqrt{2(1+\langle 1|\Sigma\rangle)}}, \tag{19}$$

and we can now plug this into our original definition of B to compute the form of this operator and of the entire T operator as a result.

3.2 Two Qubit Example

Let's now look at an example of this using just two qubits to go into more depth on the form of these operators. Let's say we want to create a Bell pair, a simple entangled state that has been realized in a number of diverse experiments [5]. This state is given as

$$|\psi\rangle = \frac{1}{\sqrt{2}}(|00\rangle + |11\rangle), \tag{20}$$

where we can see that this state is not factorable and thus entangled. In matrix notation, the statevector we want to initialize is $|\psi\rangle = \begin{pmatrix} 1/\sqrt{2} & 0 & 0 & 1/\sqrt{2} \end{pmatrix}$ (i.e. there's a 50-50 chance of measuring the pair of qubits in the $|00\rangle$ state or in the $|11\rangle$). This is a particularly powerful example of entanglement because regardless of physically separation between the two qubits, a measurement of one instantaneously determines the state of the other. This phenomenon seems so much like it violates causality that Einstein called it "spooky."

To initialize this state given in Eq. 20, we first have to find the density matrix, A, which has entries $\psi_{a,b} = \langle a, b|\psi\rangle$ from Eq. 9. If we choose the standard basis sets for $\mathcal{H}_a$ and $\mathcal{H}_b$ given by $\begin{pmatrix} 1 \\ 0 \end{pmatrix}$ and $\begin{pmatrix} 0 \\ 1 \end{pmatrix}$, then the basis set of the combined spaces is given by,

$$\text{span}(\mathcal{H}_a \otimes \mathcal{H}_b) = \{\begin{pmatrix} 1 \\ 0 \\ 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \\ 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 0 \\ 1 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 0 \\ 0 \\ 1 \end{pmatrix}\}. \tag{21}$$

Now our density matrix is trivial to compute and comes out to,

$$A = \begin{pmatrix} \frac{1}{\sqrt{2}} & 0 \\ 0 & \frac{1}{\sqrt{2}} \end{pmatrix}, \tag{22}$$

and it follows that the SVD is given by,

$$A = USV^T = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} \frac{1}{\sqrt{2}} & 0 \\ 0 & \frac{1}{\sqrt{2}} \end{pmatrix} \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}. \tag{23}$$

The next step is to compute B using Eqs. 16 & 19 using our singular values $|\Sigma\rangle = \begin{pmatrix}1/\sqrt{2} & 1/\sqrt{2}\end{pmatrix}$. Plugging this into Eq. 19 we get,

$$|q\rangle = \begin{pmatrix}0.92387953 \\ 0.38268343\end{pmatrix}, \tag{24}$$

and then plugging this into Eq. 16,

$$B = \frac{1}{\sqrt{2}}\begin{pmatrix}1 & 1 \\ 1 & -1\end{pmatrix}. \tag{25}$$

If you're familiar with quantum computing you might recognize this as the single qubit Hadamard gate, which represents a π rotation on the Bloch sphere from Fig. 2. Now we have all the components we need to construct T from Eq. 15 and use it to initialize our Bell state. The end result is the circuit seen in Fig. 4 which will properly initialize qubits $q0$ and $q1$ using the unitary operators formulated from the SVD of A.

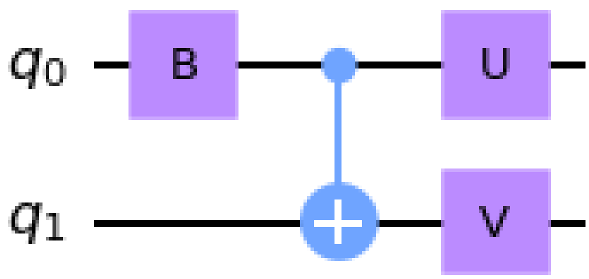

Figure 4: A quantum circuit diagram for initializing qubits $q0$ and $q1$ to a maximally entangled Bell state. In this example, B is equivalent to a Hadamard and U & V are both identity. The resulting state has a 50-50 chance of being $|00\rangle$ or $|11\rangle$.

Although we already knew that our two-qubit system was entangled from visual analysis, we can see that our Schmidt rank is greater than one which proves the entanglement of our state. Not only is our state entangled, but since our Schmidt coefficients $\lambda_1 = 1/2$, $\lambda_2 = 1/2$ are equal, we know we also have a maximally entangled state. Due to this characteristic of maximum entanglement, Bell states have common use in quantum computation systems.

The process described above for qubit state initialization can be used for arbitrary qubit states. We have shown a two-qubit system example, however systems containing many qubits are possible to initialize using the Schmidt decomposition, albeit with more complex computations. Quantum computing relies on the ability to have precise computations dependent on quantum states. Consequently, precision in state initialization is a foundational requirement for the development of quantum computing in general. The Schmidt decomposition offers a straightforward way to initialize the state of a multi-qubit system using the singular value decomposition.

References

[1] Information theory. (2020). In Encyclopædia Britannica. Retrieved from https://academic.eb.com/levels/collegiate/article/information-theory/106012

[2] Martignon, L. (2001). Information Theory. In International encyclopedia of the social & behavioral sciences / (pp. 7476–7480). Elsevier,. https://doi.org/10.1016/B0-08-043076-7/00608-2

[3] Griffiths, R.B. (2014). Hilbert Space Quantum Mechanics. Carnegie Mellon University. http://quantum.phys.cmu.edu/QCQI/qitd114.pdf

[4] Fisher, C. (n.d.). IBM — What is Quantum Computing? IBM Quantum. https://www.ibm.com/quantum-computing/learn/what-is-quantum-computing/

[5] Barnett, S. (2009). Quantum Information (Oxford Master Series in Physics, 16) (Illustrated ed.). Oxford University Press.

[6] Instituto de Física Universidade de São Paulo. (2020). Singular value and Schmidt decompisition [Slides]. IFUSP. http://www.fmt.if.usp.br/ gtlandi/04—reduced-dm-2.pdf

[7] Abhijith J., et al. Quantum Algorithm Implementations for Beginners. Los Alamos National Laboratory. https://arxiv.org/pdf/1804.03719.pdf

[8] Skript, Density Matrices, https://homepage.univie.ac.at/reinhold.bertlmann/pdfs/T2_Skript_Ch_9corr.pdf