

**University of Witwatersrand
School of Physics**

**Honours Research Report: Identifying
Quantum States using Machine
Learning
for Quantum Computation and
Communication**

By
Janice Nair

Supervisor: Dr. Isaac Nape

25 October 2023

Abstract

Decoding photonic quantum states is imperative to future quantum technologies, especially so for high-dimensional quantum systems where reconstructing the state is valuable but currently time-inefficient. An SPDC source can create pairs of photons having OAM, which may be entangled. We can take projective measurements of this bi-photon system for the purpose of reconstructing its state via Quantum State Tomography (QST). The aim of this research report was to train a Neural Network (NN) to learn to output the state given projective measurement inputs. This report discusses methods, where the most significant method had an MSE of 0.0007 for experimental data testing. This project was a 2-D demo for a future extension to high-dimensional QST.

Contents

1	Introduction	3
2	Theory	3
2.1	Required Quantum Theory	3
2.1.1	Polarization	3
2.1.2	Poincaré sphere	3
2.1.3	Density matrix	4
2.1.4	Measurement	5
2.2	Quantum State Tomography (QST)	5
3	Experiment	8
3.1	Simulating Training Data	8
3.2	Training the Neural Network	9
4	Results	10
4.1	Simulated Data Tests	10
4.2	Lab Data Tests	12
5	Conclusion	14
6	Acknowledgements	15

1 Introduction

Quantum Computing (QC), when blended with the field of Optics or Photonics, uses a photon as a Qu-bit or Quantum Bit for the purpose of information processing. Photonic Qubits have a strong significance due to their various properties that we are capable of manipulating in order to conduct information encoding, some examples include: Superposition, Polarization and Entanglement. These properties have been used in QC procedures which are important to future quantum communication networks, such examples are: Quantum Teleportation and Quantum Key Distribution (QKD).

The focus of this experiment was on the procedure of Quantum State Tomography (QST). Given some photonic quantum system, QST involves performing a series of measurements on the quantum system in different bases or measurement directions. By collecting a sufficient number of measurement outcomes, it is possible to reconstruct the full quantum state, including information about the system's superposition and entanglement properties.

The reconstruction process typically involves complex mathematical algorithms to interpret the measurement data and infer the quantum state. A significant problem in this technique is that QST can be resource-intensive and time-inefficient. Currently it is impractical for high-dimensional, large quantum systems due to the exponential growth in the number of measurements required. There are also various useful quantum systems, which require similar but unique QST techniques to decode their states. Since QST is a valuable tool, which can aid in the reliability of quantum computations, the aim of this research was to test a demo of a Neural Network trained to perform QST on a 2-dimensional, bi-photon system.

2 Theory

2.1 Required Quantum Theory

2.1.1 Polarization

Photons, as quantum particles of light, exhibit wave-particle duality. When described as waves, as in Figure 1, photons have an associated electric field. The polarization of a photon emerges as a property related to the orientation of this electric field vector.

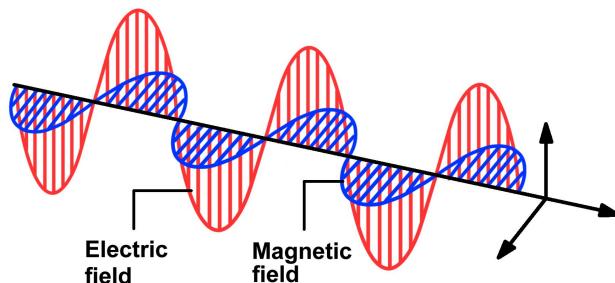


Figure 1: Electromagnetic wave[1].

The “wavefunction” however in the context of quantum mechanics describes a particle’s quantum state. And a photon can exist in a superposition of polarization states, such as horizontal, vertical, diagonal, or circular polarization.

2.1.2 Poincaré sphere

The Poincaré Sphere is a 3-D coordinate system used to represent the polarization states of a photon in the form of vectors. In this report the words “vector” and “state” are used synonymously. The reason for this is because the state of a photon can be mapped onto the Poincaré sphere, shown in Figure 2 (c). As such, basis states and photon states would also be vectors themselves.

Key points about the Poincaré sphere:

1. 3D Sphere: It uses spherical coordinates, where any vector point on or in the sphere corresponds to a unique state for the photon.
2. The basis vectors chosen are shown in Table 1.

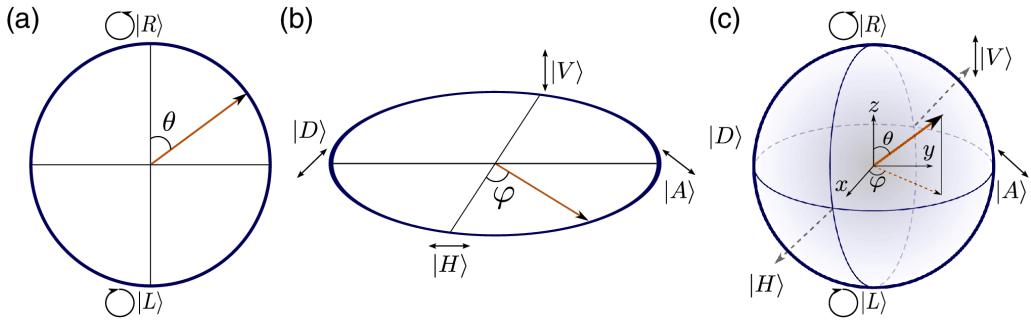


Figure 2: (a) \$z - y\$ plane and (b) \$x - y\$ plane of (c) The Poincaré Sphere[3].

3. Linear combinations of states.

Points on the Sphere; represent “pure states”.

Points Inside the Sphere; represent “mixed states”. These states have an equal mixture of two or more pure polarization states[4].

The south-pole of the sphere represents Left-circular polarization \$|L\rangle\$, the north-pole represents Right-circular polarization \$|R\rangle\$, and the equator represents linear-polarization (including Horizontal, Vertical, Diagonal and Anti-diagonal). In summary, the Poincaré sphere provides a comprehensive representation of photon states or qubit states.

n	Polarization Basis Vector	Description	Phases	$\begin{matrix} +\pi \\ -\pi \end{matrix}$
1	\$ R\rangle = \begin{pmatrix} 1 \\ 0 \end{pmatrix}\$	Right circularly polarized		
2	\$ L\rangle = \begin{pmatrix} 0 \\ 1 \end{pmatrix}\$	Left circularly polarized		
3	\$ H\rangle = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 \\ i \end{pmatrix}\$	Horizontally polarized		
4	\$ V\rangle = \frac{-i}{\sqrt{2}} \begin{pmatrix} 1 \\ -i \end{pmatrix}\$	Vertically polarized		
5	\$ D\rangle = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 \\ 1 \end{pmatrix}\$	Diagonally polarized		
6	\$ A\rangle = \frac{i}{\sqrt{2}} \begin{pmatrix} 1 \\ -1 \end{pmatrix}\$	Anti-diagonally polarized		

Table 1: Polarization Basis Vectors of Poincaré Sphere.

2.1.3 Density matrix

A single photon in a pure state, eg. \$|\Psi\rangle = \tilde{\alpha}|R\rangle + \tilde{\beta}|L\rangle\$, can have its state represented as:

$$|\Psi\rangle = [\tilde{\alpha}, \quad \tilde{\beta}] . \quad (1)$$

Where \$\tilde{\alpha}\$ and \$\tilde{\beta}\$ are complex probability amplitudes. This state is also referred to as its wavefunction. It also has a density matrix given by the outer product:

$$\vec{\rho}_{pure} = |\Psi\rangle \langle \Psi| = \vec{\rho}, \quad (2)$$

Which has the commonly known normalization,

$$tr(\rho_{pure}^2) = \int_{-\infty}^{\infty} |\Psi(x)|^2 dx = 1, \quad (3)$$

Where \$\text{tr}()\$ is the trace of the matrix, and \$x\$ is the spatial coordinates of the wavefunction (or state) \$\Psi(x)\$. However, mixed states are ensembles of pure states. A mixed state arises when a quantum system is not in a well-defined pure state but is instead described by a probabilistic linear-combination of different pure states,

each with associated probabilities C_m . A mixed state photon $|\psi\rangle_{mixed}$ has its density matrix given by the outer product:

$$\vec{\rho}_{mixed} = \sum_{m=1}^N C_m |\Psi_m\rangle \langle \Psi_m|, \quad (4)$$

With the conditions:

$$0 \leq \text{tr}(\vec{\rho}_{mixed}) < 1, \quad \sum_{m=1}^N C_m = 1, \quad (5)$$

Where N is the total number of pure states that $\vec{\rho}_{mixed}$ is dependent on. For this report we only focus on pure state QST.

2.1.4 Measurement

Measurement of a quantum state can also be referred to as “projective measurement” in our case. This is because the basic concept is that: we can project some photon in question onto each of the 6 polarization basis vectors defined previously and yield an expectation value or “measurement” which tells us how much the photon-state belongs to each basis vector.

Mathematically, to make a matrix of projective measurements \vec{M}_n , for the set of $n = [1, 6]$ polarization bases $P_n = (R, L, H, V, D, A)$, you would use the following expressions:

$$\vec{M}_n = [m_1, \ m_2, \ m_3, \ m_4, \ m_5, \ m_6], \quad (6)$$

$$m_n = \langle P_n | \vec{\rho} | P_n \rangle = \vec{\rho} \times \vec{Proj}_n, \quad (7)$$

Where $\vec{\rho}$ is the density matrix of some photon state, and $\vec{Proj}_n = |P_n\rangle \langle P_n|$ is the projector of each of the six polarization bases. The expression $\langle P_n | \vec{\rho} | P_n \rangle$ is simply the expectation value of the state with respect to polarization basis P_n , also synonymous with the “measurement” of the photon-state in P_n polarization. An example of a measurement matrix (normalized to 1) for the single photon is visualized in Figure 3 below. The higher that a normalized measurement value is, the more photons were found at that polarization.

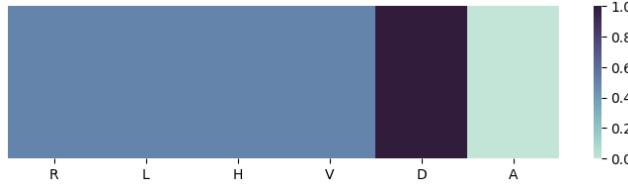


Figure 3: Heatmap of a measurement matrix \vec{M}_n , which corresponds to the single-photon state $|\Psi\rangle = \left[\frac{1}{\sqrt{2}}, \ \frac{1}{\sqrt{2}} \right]$

2.2 Quantum State Tomography (QST)

Tomography of a quantum state is the process of reconstructing the state of a quantum system to identify it. A laser beam emits a total amount of N photons. If the laser emits pure state photons this means that all the photons emitted are at the same state $|\psi\rangle$. But, if the source emits mixed state photons then each emitted photon can be described as having a unique state $|\psi_m\rangle$. The procedure of finding the state relies on data acquisition in the form of counts of photons N_{P_n} - this is the amount of photons measured at basis P_n , where $\sum_1^n N_{P_n}$. We can use this to determine a matrix of ratios:

$$\left[\frac{\{N_{P_n}\}}{\max(N_{P_n})} \right] = \vec{M}_n, \quad (8)$$

Where \vec{M}_n is the projective measurement matrix discussed previously. When conducting a QST the systems \vec{M}_n would be measured and then used to reconstruct the state.

QST of a Bi-Photon System (2 Photons)

A Gaussian Beam laser pump source, with topological charge $\ell = 0$, generates a single photon. Then inputting the laser source into a non-linear crystal generates a bi-photon system via Spontaneous Parametric Down Conversion (SPDC). One of the output photons has $\ell = 1$ and the other photon has $\ell = -1$, this is shown in Figure 4. Hence, these photons each have an OAM state $|\ell\rangle$ and $|-\ell\rangle$ which can be interpreted as the right and left circular photon polarization, respectively.

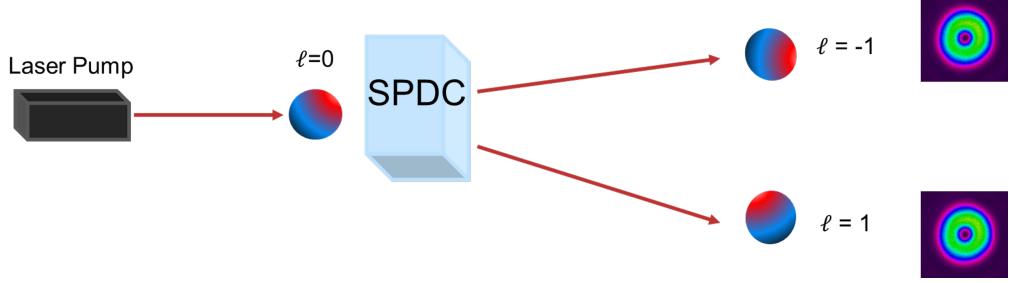


Figure 4: Experimental generation of the bi-photon state $|\Psi\rangle = \left[\frac{1}{\sqrt{2}}, \quad \frac{1}{\sqrt{2}} \right]$.

This bi-photon system may often generate entangled pairs. A multi-photon (or qubit) state, is a tensor product of the wavefunctions of each of the photons. If we have a photon A and a photon B, the joint state is:

$$|\Psi_{AB}\rangle = |\Psi_A\rangle \otimes |\Psi_B\rangle = (\tilde{\alpha}|\ell\rangle + \tilde{\beta}|-\ell\rangle) \otimes (\tilde{\alpha}|\ell\rangle + \tilde{\beta}|-\ell\rangle) \quad (9)$$

$$= \tilde{a}|\ell\rangle_A|\ell\rangle_B + \tilde{b}|\ell\rangle_A|-\ell\rangle_B + \tilde{c}|-\ell\rangle_A|\ell\rangle_B + \tilde{d}|-\ell\rangle_A|-\ell\rangle_B, \quad (10)$$

Where the co-efficients of the tensor product multiply with each other like: $\tilde{a} = \tilde{\alpha}\tilde{\alpha}$, $\tilde{b} = \tilde{\alpha}\tilde{\beta}$, $\tilde{c} = \tilde{\beta}\tilde{\alpha}$ and $\tilde{d} = \tilde{\beta}\tilde{\beta}$. Using the polarization vectors defined in Table 1, the state can then be written as:

$$|\Psi_{AB}\rangle = \tilde{a} \begin{pmatrix} 1 \\ 0 \\ 0 \\ 0 \end{pmatrix} + \tilde{a} \begin{pmatrix} 1 \\ 0 \\ 0 \\ 0 \end{pmatrix} + \tilde{c} \begin{pmatrix} 0 \\ 0 \\ 1 \\ 0 \end{pmatrix} + \tilde{d} \begin{pmatrix} 0 \\ 0 \\ 0 \\ 1 \end{pmatrix}, \quad (11)$$

$$|\Psi_{AB}\rangle = \begin{pmatrix} \tilde{a} \\ \tilde{b} \\ \tilde{c} \\ \tilde{d} \end{pmatrix}. \quad (12)$$

Since each photon has 6 possible polarization states, we will then also have $6 \times 6 = 36$ combinations of projectors for the joint state $|\Psi_{AB}\rangle$. Which will result in 36 projective measurements to be taken in order to determine the co-efficients of the state $(\tilde{a}, \tilde{b}, \tilde{c}, \tilde{d})$. This means that \vec{M}_n would have 36 elements, as seen in the example visualized in Figure 5(a).

$$\vec{M}_n = \begin{bmatrix} m_1, & m_2, & m_3, & m_4, & m_5, & m_6 \\ m_7, & m_8, & m_9, & m_{10}, & m_{11}, & m_{12} \\ m_{13}, & m_{14}, & m_{15}, & m_{16}, & m_{17}, & m_{18} \\ m_{19}, & m_{20}, & m_{21}, & m_{22}, & m_{23}, & m_{24} \\ m_{25}, & m_{26}, & m_{27}, & m_{28}, & m_{29}, & m_{30} \\ m_{31}, & m_{32}, & m_{33}, & m_{34}, & m_{35}, & m_{36} \end{bmatrix}, \quad (13)$$

$$m_n = \langle P_n | \vec{\rho} | P_n \rangle = \vec{\rho} \times \vec{Proj}_n. \quad (14)$$

The bi-photon state is a 2-dimensional state having 2 basis states that are used to span the state space; $|\ell\rangle$ and $|-\ell\rangle$. This means that amount of measurements we have to take are dependent on the equation $(d^2 + d)^n$ where the dimension of the photon state is d and the number of photons in the state is n . So it is worth noting that as you increase the dimensionality of a quantum photonic systems, the amount of measurements you have to take scales exponentially-this fact makes the QST process time consuming, computationally intensive and storage heavy. Solving those problems is important to this research report and is the goal for future studies. Hence the bi-photon system, in Figure 6, was used as a demo for this report.

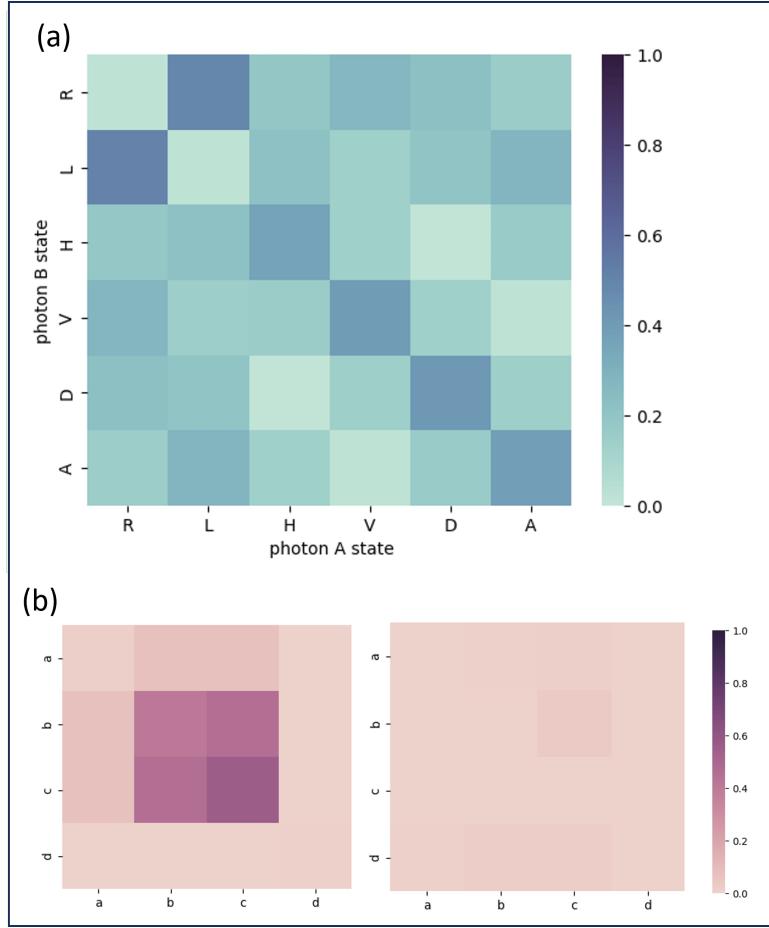


Figure 5: Example: (a) Experimentally obtained measurement matrix of the state $\begin{bmatrix} 0 & \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} & 0 \end{bmatrix}$.(b) Corresponding density matrix found by brute force method.

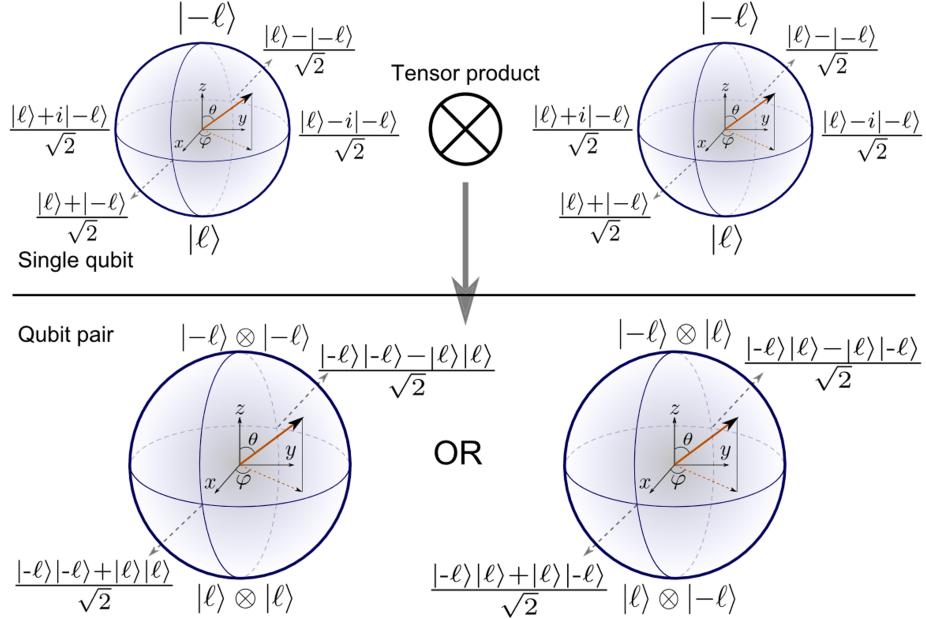


Figure 6: Visualization of the bi-photon joint state. Where superpositions of $|\ell\rangle$ and $|-\ell\rangle$ are the new polarization states.

3 Experiment

The experiment was to train a neural network to predict the bi-photon quantum state if it's trained on projective measurements. The method for such problems usually follows: (i) Simulating Training Data (ii) Training the Neural Network (iii) Testing the Neural Network. We first simulated the state space, then calculated the density matrix through Equation 2 for each state in the space, we then found the measurement matrix via Equations 13 and 14. The states made up the output data, and the measurements made up the input data, of the Neural Network as seen in Figure 7.

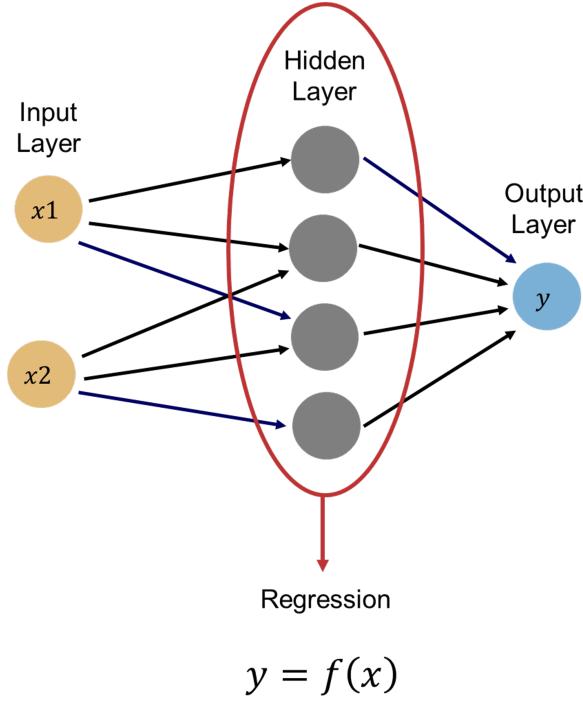


Figure 7: Simplified Neural Network Graphic. Trained to learn the linear regression function $f(x)$ -this function represents QST.

3.1 Simulating Training Data

Generating States to Span the State-Space

The process of data simulation was carried out in Python and stored in pickle files to be called for training the model in TensorFlow with Keras. The data simulation and model training code can all be found in the github link[2]. The word span refers to how comprehensive the data is so that it can model the state space of the bi-photon and generalize it well. As the starting point, the aim was to train a model which predicts the complex probability amplitudes of a state:

$$|\Psi_{AB}\rangle = \begin{pmatrix} \tilde{a} \\ \tilde{b} \\ \tilde{c} \\ \tilde{d} \end{pmatrix}. \quad (15)$$

We treated this as a problem which required spanning the state space of the bi-photon system. The requirements were that:

$$|\tilde{a}|^2 + |\tilde{b}|^2 + |\tilde{c}|^2 + |\tilde{d}|^2 = 1. \quad (16)$$

This is a commonly known normalization of the probability amplitudes of pure states. Another constraint on individual probability amplitudes is that, based on Equation 16, each must have real and complex parts within the range of $[-1, 1]$. In Python, we populated 5 000 000 size 8 arrays using a random uniform number generator from NumPy. The array size of 8 was chosen to account for 4 probability amplitudes per state, each with real and imaginary parts. This resulted in 8 significant elements, or features in this case. This choice also removed the imaginary number i from the state arrays, because the TensorFlow neural network excludes complex numbers when training. The uniform number generator gave the datapoints in the range of $[-1, 1]$. This random generation is visualized as a data distribution in Figure 8, where a 2000 data-point sample was used

for visualization purposes. But to make the state arrays adhere to the constraint in Equation 16 we multiplied each array by the quantum normalization factor:

$$\frac{1}{\sqrt{|\tilde{a}|^2 + |\tilde{b}|^2 + |\tilde{c}|^2 + |\tilde{d}|^2}}. \quad (17)$$

The resulting data distribution was closer to spanning the state space, and can be seen visualized in Figure 9.

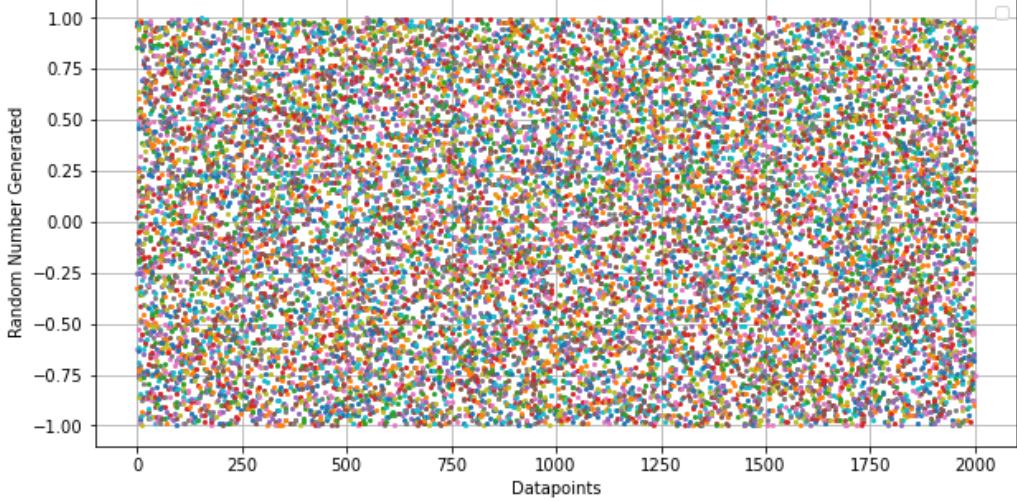


Figure 8: Random Uniform Data Distribution.

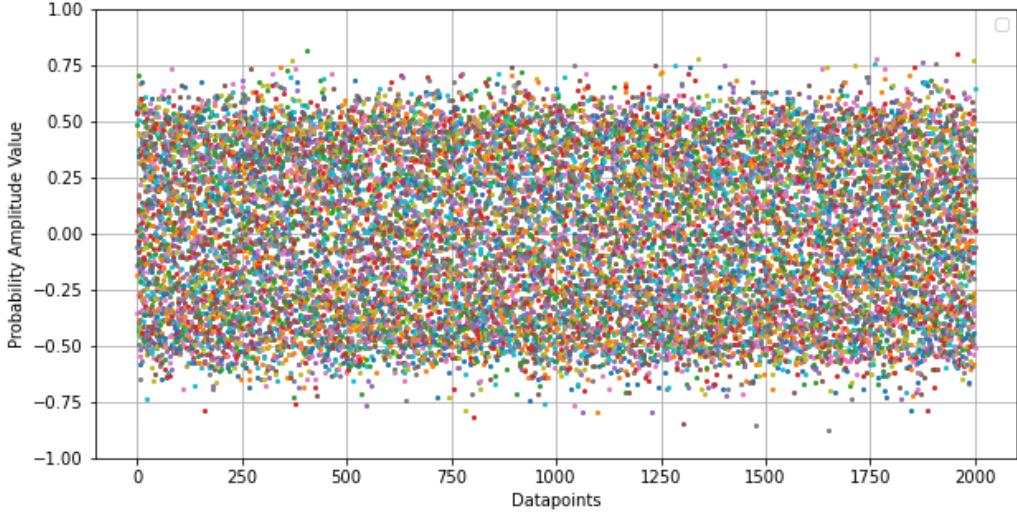


Figure 9: Data Distribution of Normalized Probability Amplitudes.

3.2 Training the Neural Network

The training process heavily depended on hyperparameter tuning, parameters such as batch size and epoch size were optimized according to the data distribution shown previously. The first neural network that was trained, a simple 3-layer network with LeakyReLU activation, was trained on a limited span of the probability amplitudes- [0, 1], and was highly effective at predicting unique random states. For the second attempt at the model training a span of [-1, 1] was used instead, this resulted in huge errors and early convergence. The cause being likely a vanishing gradient descent for large batch sizes and an exploding gradient for small batch sizes. This hinted that the complexity of the function we were trying to model was too high for the structure of the neural network as it was. Hyperparameter tuning managed to find a minima of a 0.0011 MSE loss, this is highly unlikely to be the true global minimum of this problem, but it was nevertheless tested on simulated data and experimentally obtained data. This is the main model that the Results section details.

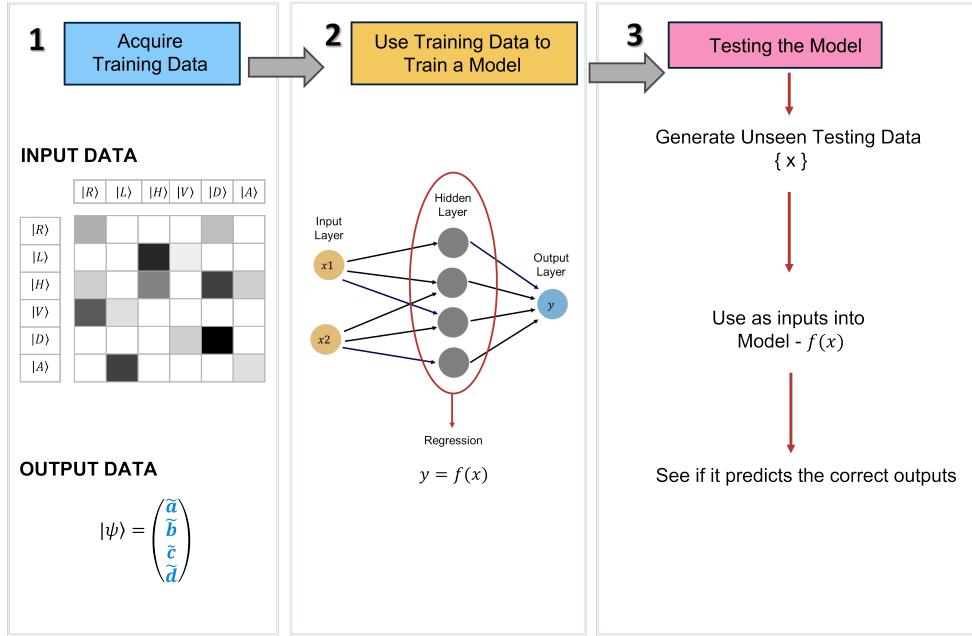


Figure 10: The process for training the neural network, with a summary of the training features.

4 Results

The interpretation of the results is based off testing the models predictions of various entangled and non-entangled, noisy state measurements. The example of lab data shown in Figure 5 was predicted by the model and displayed in Figure 11 as density matrices.

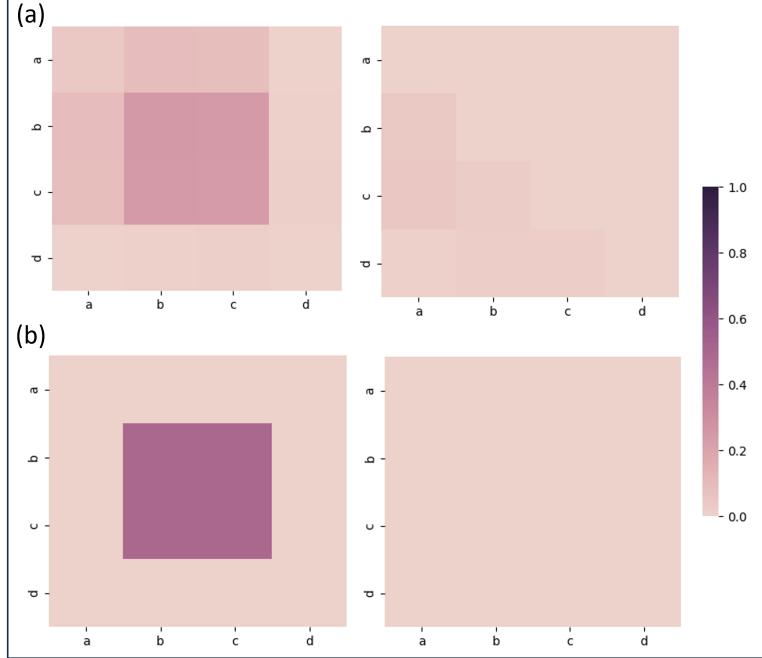


Figure 11: Experimental Data: (a) Predicted density matrix compared to (b) theoretical density matrix for the state $\left[0 \quad \frac{1}{\sqrt{2}}, \quad \frac{1}{\sqrt{2}} \quad 0\right]$.

4.1 Simulated Data Tests

To first test the capabilities of the model, tests on unique measurement data simulated the same way as in the Experiment section were used to determine an average error in Figure 12. Some examples of states that show the accuracy of these predictions are shown in Figure 13.

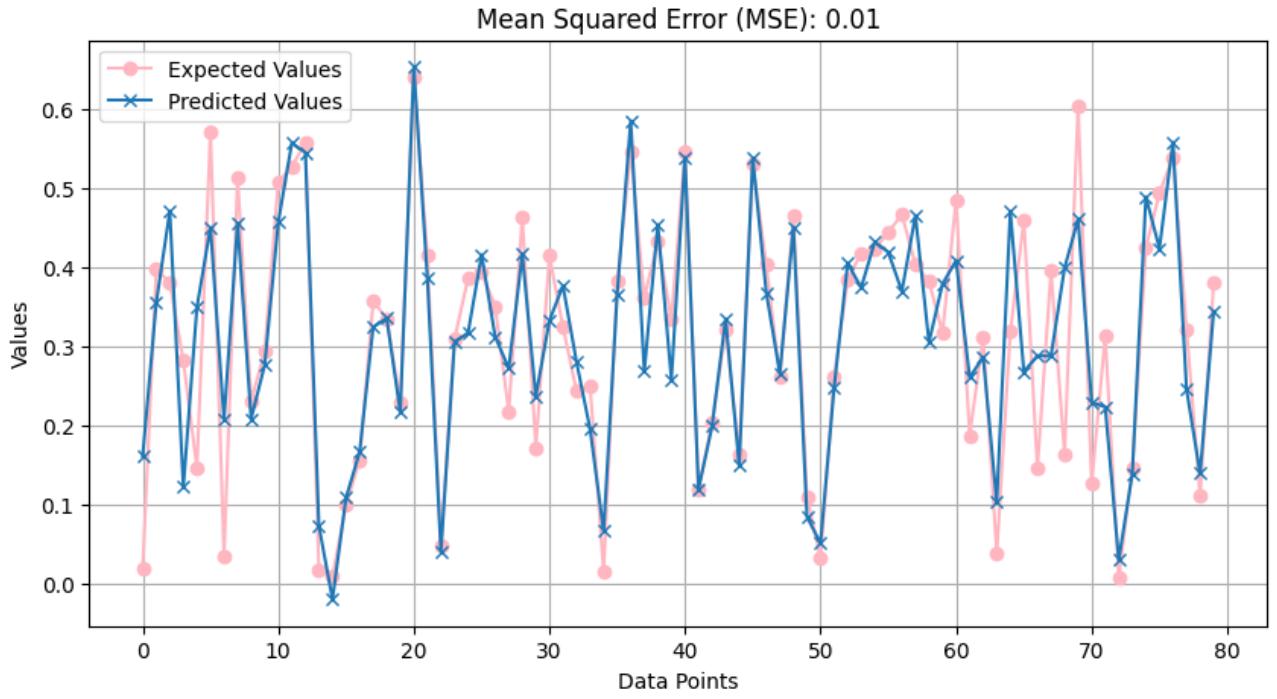


Figure 12: Mean Squared Error, with average error of 5% found for randomly generated states and the models predictions of them.

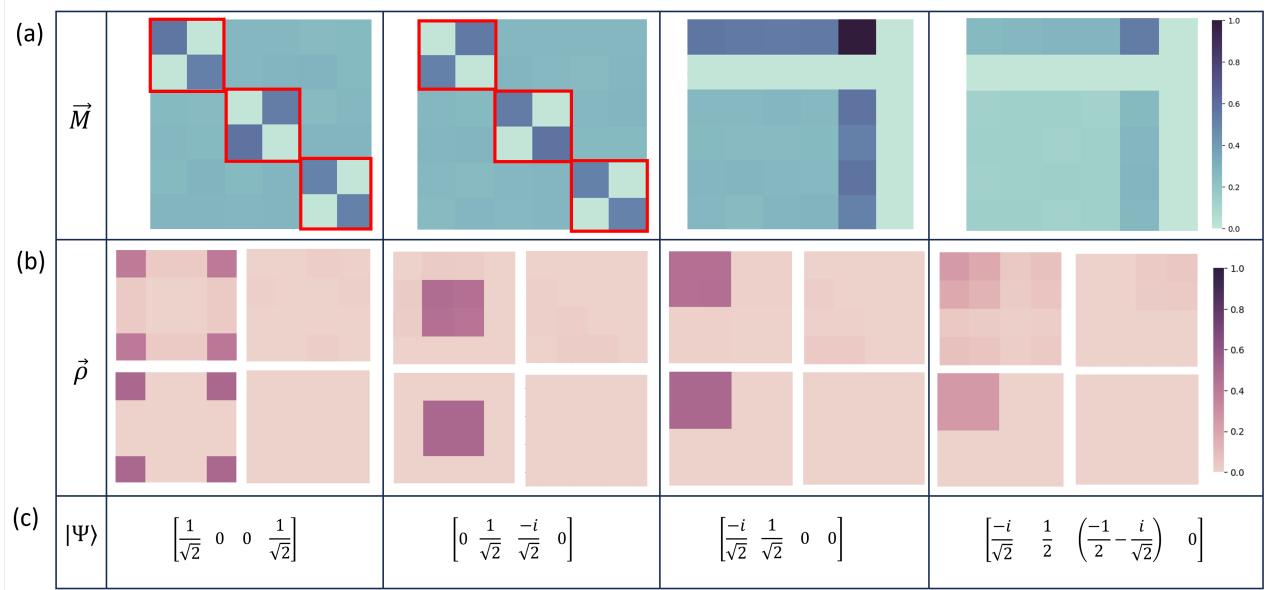


Figure 13: Examples of Simulated Data Predictions: (a) Noisy measurement matrices \vec{M} of the sample states $|\Psi\rangle$. (b) Top: Predicted States, real and imaginary plots, displayed as density matrices. Bottom: Theoretical density matrices.

4.2 Lab Data Tests

The main purpose of this research being for the NN model to predict real-world noisy data, these lab tests were done on data experimentally obtained. These measurements were taken from an experiment as the one described in Section 2.2. The set of examples used are visualized in Figures 14 and 15. The accuracy of these predictions is analysed in Figure 1 where an average accuracy of 4.166% was determined for these predictions.

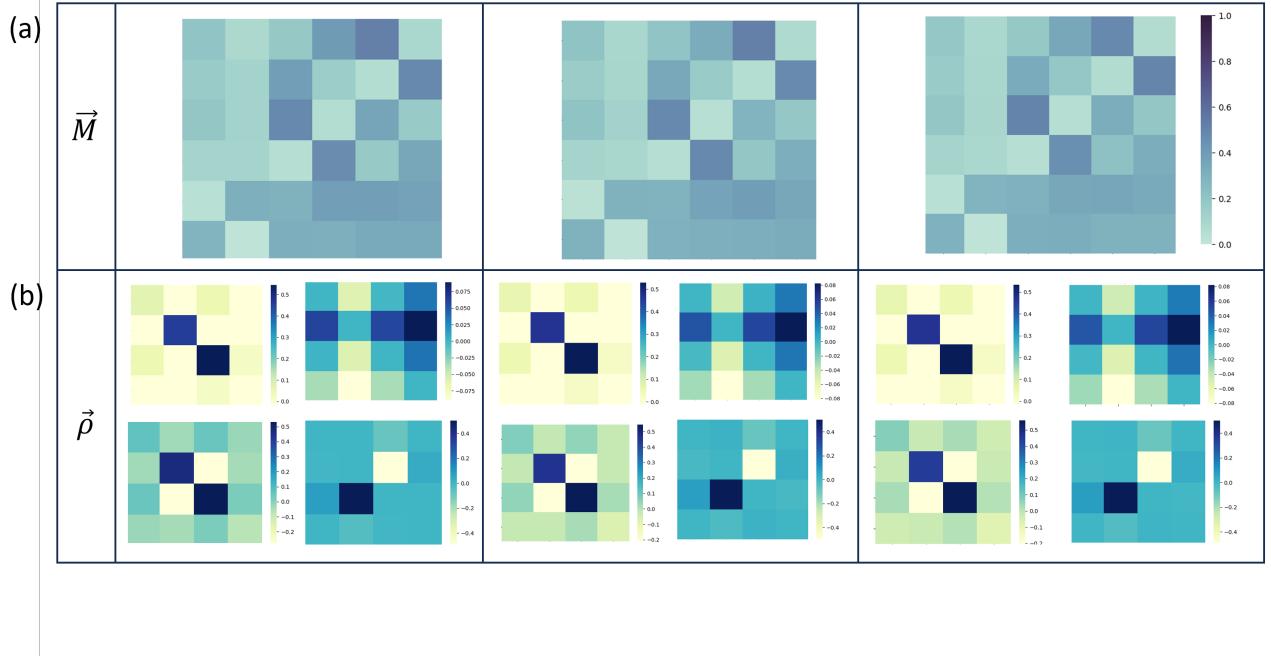


Figure 14: Experimental Data Predictions: (a) Experimentally obtained measurement matrices \vec{M} of the sample states $|\Psi\rangle$. (b) Top: Predicted states, the real and imaginary plots of the density matrices. Bottom: Theoretical density matrices.

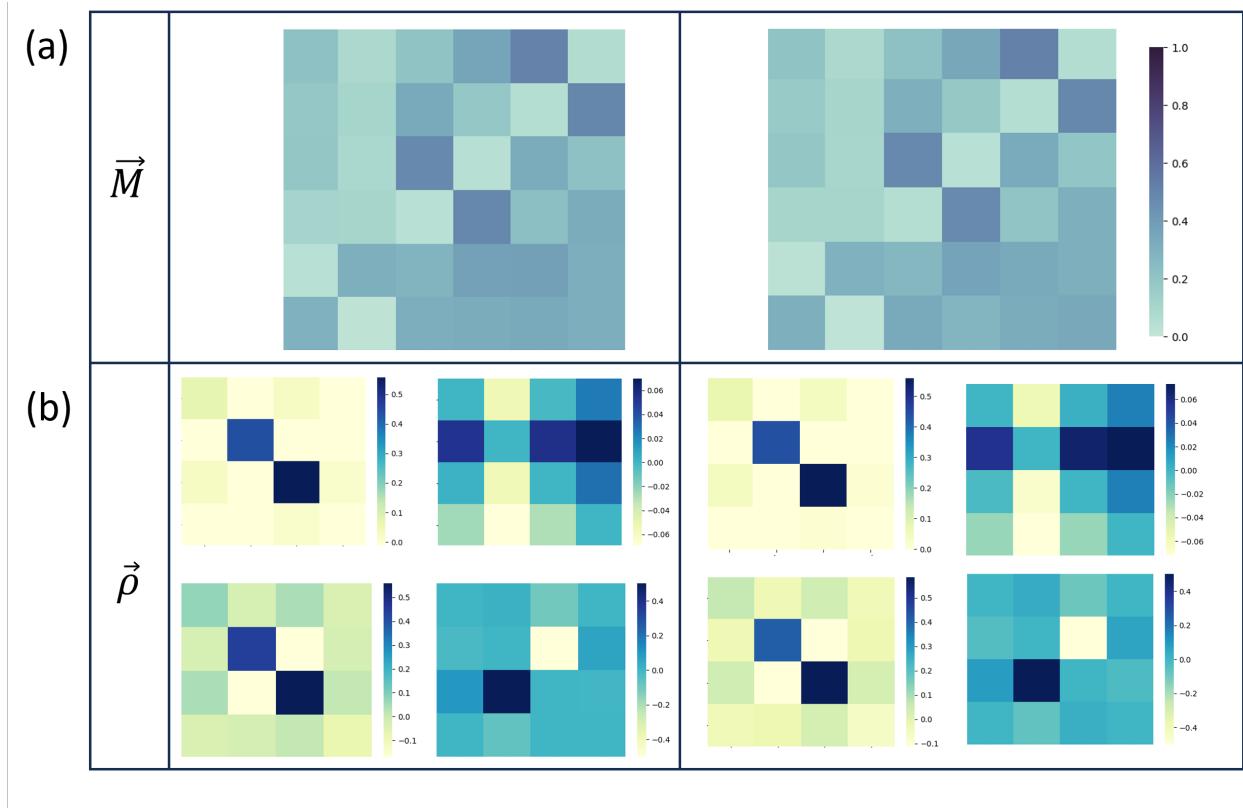


Figure 15: Experimental Data Predictions: (a) Experimentally obtained measurement matrices \vec{M} of the sample states $|\Psi\rangle$. (b) Top: Predicted states, the real and imaginary plots of the density matrices. Bottom: Theoretical density matrices.

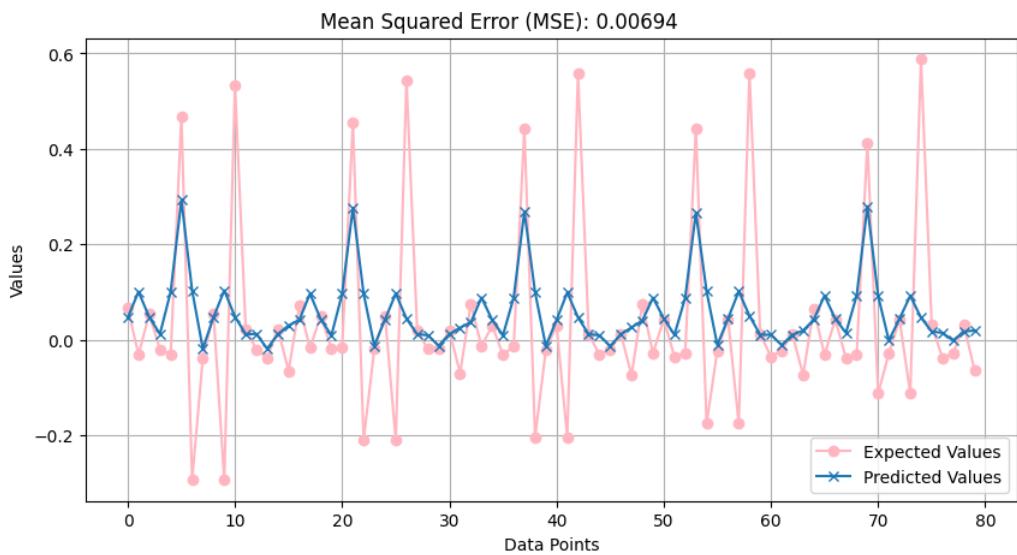


Figure 16: Mean Squared Error of predicted density matrices against the corresponding experimental density matrices. This was a RMSE:0.0833, or an average error of 4.166%.

5 Conclusion

The results are promising that a neural network can perform QST and predict high dimensional states, an RMSE of 0.0833 or a percentage error of 4.165% was observed for tests on experimental data. However, the subset of data being tested is still small and requires a model that is well generalized in order to identify varieties of states. The training is inadequate due to the gradient vanishing issues experienced. There are various avenues to explore regarding this problem, in terms of model tuning and data generation. The model can be improved to match the complexity of the data, or the data can be generated functionally to match the problem space in more detail. This would include using the equations of the state at various polarization's and this could give more comprehensive information to the model than random state generation. A wide range of applications exist which often involve substantial data processing in quantum communication channel characterization and quantum imaging, particularly in the context of image reconstruction. Integration of experiments with tools such as this NN and other optimization and data processing architectures, holds the potential to yield practical advantages for technologies and introduces further novel research opportunities due to the improvements in time-efficient data acquisition.

6 Acknowledgements

Much appreciation to everyone in the Structured Light Lab who aided this research project tremendously, in particular: my supervisor Isaac Nape, as well as Neelan Goudan and Pedro Ornelas.

References

- [1] Mr Kendrick and Seyed Hosseini. Wave particle and luminiferous aether trinity of the light introduction. 02 2018.
- [2] Janice Nair. Identifying-quantum-states-with-machine-learning, 2023. <https://github.com/janybear/Identifying-Quantum-states-with-Machine-Learning> (Accessed: October 25, 2023).
- [3] Ermes Toninelli, Isaac Nape, Andrew Forbes, and et al. Concepts in quantum state tomography and classical implementation with intense light: a tutorial. *Adv. Opt. Photon.* 11, 67-134, Vol. 11(Issue 1):73, 2019.
- [4] Professor Steven van Enk. Mixed states and pure states. Course Notes, April 9, 2009. PHYS 415-Quantum Mechanics II (undergrad level), University of Oregon.