Artificial Neural Networks in Quantum Information and Nuclear Magnetic Resonance

Ankit Kumar

A dissertation submitted for the partial fulfilment of BS-MS dual degree in Science

Under the guidance of

Prof. Kavita Dorai

Indian Institute of Science Education and Research Mohali

May 2020

Certificate of Examination

This is to certify that the dissertation titled **"ANN in Quantum Information and NMR"** submitted by **Ankit Kumar** (Reg. No. MS15009) for the partial fulfillment of BS-MS dual degree programme of the Institute, has been examined by the thesis committee duly appointed by the Institute. The committee finds the work done by the candidate satisfactory and recommends that the report be accepted.

Dr. M. Bera	Prof. Arvind	Prof. Kavita Dorai
		(Supervisor)

Dated: 31.05.2020

Declaration

The work presented in this dissertation has been carried out by me under the guidance of Prof. Kavita Dorai at the Indian Institute of Science Education and Research Mohali.

This work has not been submitted in part or in full for a degree, a diploma, or a fellowship to any other university or institute. Whenever contributions of others are involved, every effort is made to indicate this clearly, with due acknowledgement of collaborative research and discussions. This thesis is a bonafide record of original work done by me and all sources listed within have been detailed in the bibliography.

Ankit Kumar (Candidate)

Dated: May 31, 2019

In my capacity as the supervisor of the candidate's project work, I certify that the above statements by the candidate are true to the best of my knowledge.

Prof. Kavita Dorai
(Supervisor)

Acknowledgement

First and foremost, I would like to express my gratitude to my thesis supervisor Prof. Kavita Dorai, without whose help and supervision, this thesis would have never been possible. I would also like to thank him for giving me space and freedom to explore the subject. The discussions that I had with him has enhanced my capabilities as a researcher.

I also thank Prof. Arvind and Dr. M. Bera for being my committee members. The discussions I had with Akshay Gaikwad and other members of our lab helped me to gain an in-depth understanding of the subject. I would like to thank them as well.

I'm grateful to IISc Bangalore, for providing me financial support through KVPY grants. And above all, I whole-heartedly thank IISER Mohali for providing me the work-space and a sound environment to pursue my research.

Finally, thankful to all of my friends especially Satyam P., Himanshu, Pratibha, and, to my brother and my parents for their love and support, throughout my entire journey.

Ankit Kumar

MS15009

IISER Mohali.

List of Figures

Notations

$\|0\rangle$	$\begin{pmatrix} 1 \\ 0 \end{pmatrix}$
$\|1\rangle$	$\begin{pmatrix} 0 \\ 1 \end{pmatrix}$
$\|+\rangle$	$\begin{pmatrix} 1 \\ 1 \end{pmatrix}$
$\|-\rangle$	$\begin{pmatrix} 1 \\ -1 \end{pmatrix}$
σ_x or X	$\begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$
σ_y or Y	$\begin{pmatrix} 0 & -\iota \\ \iota & 0 \end{pmatrix}$
σ_z or Z	$\begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}$
H (Hadamard matrix)	$\begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix}$
$\mathbb{F}_2^n$	n-tuple space with entries $\in \{0, 1\}$
$M^{\otimes n}$	n-tensor product of M

Contents

Abstract

Artificial neural networks(ANN) imitated to biological neural networks constituting network of neuron which learns from data and the computing systems. Machine Learning(ML) is a subset of Artificial Intelligence(AI), which learns from data, examples, and without being explicitly programmed. A variety of application has found of ANN in Qunatum Information like Entanglement Detection of Quantum System, study NMR(Nuclear Magnetic Resonance) spectra.

One can clasiify Artificial neural networks into discrete-variable and continuous-variable artificial neural network. A comparison of efficiency has been made between these two networks with their cost.

The PPT(Partial positive transpose) criterion uses to detect entanglement for bipartite quantum systems, here we use ANN model and PPT criteria for qubit-qubit entanglement detection, and Entanglement criteria for Qutrits.

ANN enables quantification of spectra got from NMR, like structure elucidation, peak, phase shift. Analyses Lineshift fitting and does lipoprotein isolation by density of protein through ANN.

Chapter 1

Introduction

Artificial neural networks(ANN) learns from data and computing systems, also said to be methodologies of ML(Machine Learning). In early 1940s, it started but the growth and power of computational gained in late 20th century, when it has got high-level representation by using successive layers of real-valued latent variables or binary with a Boltzmann machine. Where unsupervised pre-training increased a computing power and also from GPUs to allow the use of big networks.

The neural networks has been applied in different field and it's appreciation can seen in Quantum Mechanics also like detecting entanglement of quantum states, Phase shift, peak detection of NMR Spectra, and also isolation of lipoprotein via density of protein through ANN.

As discrete-variable network has unitary operators of hidden layers, which acts on input sequentially with positive transpose map, while continuous have infinite dimensions Hilbert space.

The PPT(Partial positive transpose) criterion can uses to detect entanglement for bipartite quantum systems, here we use ANN model and PPT criteria for qubit-qubit entanglement detection, and entanglement criteria for Qutrits, where qutrits were in a mixture of both individually the maximally mixed state and a maximally entangled state.

ANN enables quantification of spectra got from NMR, like structure elucidation, peak, phase shift.

We look entanglement for two qutrits system which are a mixture of both individually the maximally mixed state and a maximally entangled state, where we derived that probability should less than 1/4 for being separable.

Most of the time we keep testing data set much larger than validation data sets, to get much better learning, and if validation loss is more compare to testing, we can say, we have not defined enough hidden layers, or bias function, because of that we can not have efficient models of artificial neural networks.

Chapter 2

Artificial Neural Networks

We will look the training of neural networks with Forward Propagation, Backward Propagation with partial derivatives and Loss Evaluation, then about discrete-variable neural network and continuous-variable neural networks.

2.1 Training the network

ANN is trained by using Perceptron, which has basic units of neurons, and associated with learning rule.

Learning Steps: Randomization of parameter, Forward Propagation(Linear Combination, Activation), Backward Propagation(partial derivatives, Loss Evaluation), Update parameter. They try to minimize cost functions by using Gradient Descent rule.

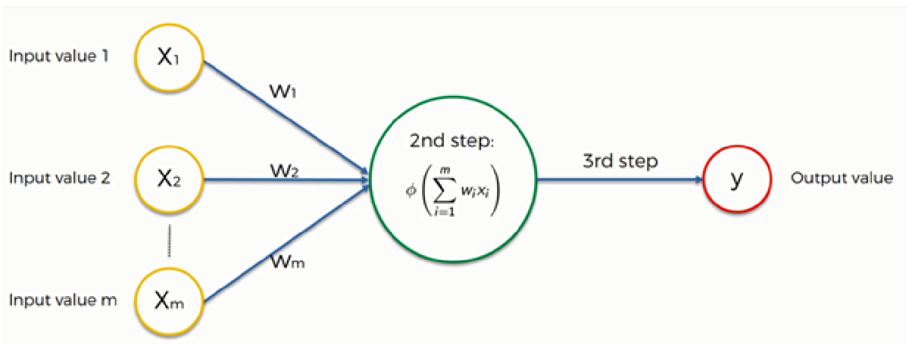

Figure 2.1: The schematic diagram of the Neural Network

- X1, X2 and Xm are the inputs
- W1, W2 and Wm are the corresponding weights

• For b to be the perceptron's bias

• y is the output

The output of the perceptron y as, y=f(W1* X1+ W2* X2+...+ Wm* Xm+ b); A full connection of network can be seen with linear activation function and sigmoid activation function.

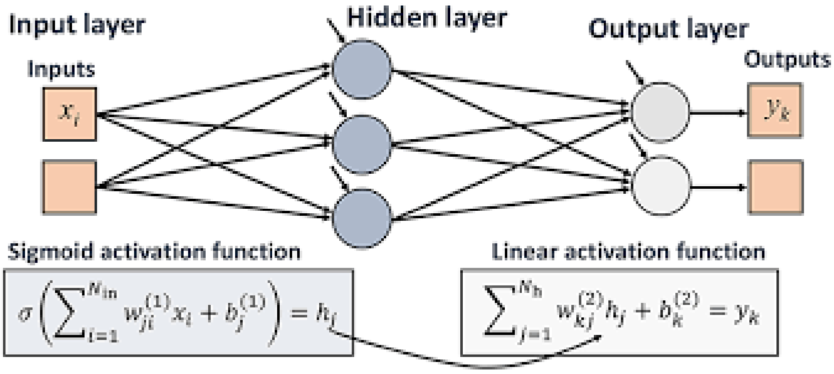

Figure 2.2: Activation function and output of Network

Forward Propagation: The Forward Propagation needs the input (x) being fed to the network, the activation function, and the weight function associated with the neurons. The evaluation and storage of outputs includes the intermediate variables for the neural network, and the network continues to propagates forward to generate its output say(y), if $X_{a \times b}$ be the input matrix, and the output $Y_{c \times b}$. Then for the ith layer:

$$Z^i_{N_i \mathrm{xb}} = W^{i-1}_{N_i \times \mathrm{a}} * A^{i-1}_{a \times b} + B^i_{N_i \times b}$$
$$A^i_{N_i \times b} = f\left(Z^i_{N_i \times b}\right)$$

The weight matrix between layers $i-1$ and i is $W^{i-1}_{N_i \times a}$ and $B^i_{N_i \times b}$ is the ith layer bias function. We can use any of two activation function accordingly,

Two activation function: $f\left(Z^i\right) = \begin{cases} \tanh\left(Z^i\right) & i \neq N_f \\ \operatorname{sigmoid}\left(Z^i\right) = \frac{1}{1+e^{-Z^i}} & i = N_f \end{cases}$

Backward Propagation: Once we evaluated the output, we can go back to check how it has correlated with estimated output with partial derivatives, and loss Evaluation or function, once we do the backward propagation then update the parameters accordingly,

Say $\hat{Y}_{c \times b}$ to evaluated output and $Y_{c \times b}$ to the know output, then error function can be write as,

$$E = \mathcal{L}\left(\hat{Y}_{c \times b}, Y_{c \times b}\right) \tag{2.1}$$

and the partial derivative can evaluate by chain rule and taking derivative of loss function w.r.t. parameters A,[Wittek 14]

For the ith layer:

$$dA = \frac{\partial \mathcal{L}}{\partial A} \tag{2.2}$$

$$dZ = \frac{\partial \mathcal{L}}{\partial Z} = \frac{\partial \mathcal{L}}{\partial A} * \frac{\partial A}{\partial Z} = dA * \frac{\partial A}{\partial Z} \tag{2.3}$$

$$dW = \frac{\partial \mathcal{L}}{\partial A} * \frac{\partial A}{\partial Z} * \frac{\partial Z}{\partial W} \tag{2.4}$$

For the $(i-1)$th layer:

$$dA^{i-1} = \frac{\partial \mathcal{L}}{\partial A^{i-1}} = \frac{\partial \mathcal{L}}{\partial A^{i}} * \frac{\partial A^{i}}{\partial Z^{i}} * \frac{\partial Z^{i}}{\partial A^{i-1}} \tag{2.5}$$

$$= dZ^{i} * W^{i-1} \tag{2.6}$$

So,

$$dZ^{i-1} = dA^{i-1} * \frac{\partial A^{i-1}}{\partial Z^{i-1}} \tag{2.7}$$

$$dW^{i-2} = dZ^{i-1} * \frac{\partial Z^{i-1}}{\partial W^{i-2}} \tag{2.8}$$

Similarly we can apply this for the Bias functions.

Then we can minimize the cost function, which is one half of the mean squared error, from gradient descent method,

$$J\left(\theta_0, \theta_1\right) = \frac{1}{2N} \sum_{i=1}^{N} \left(f_\theta\left(x^{(i)}\right) - y^{(i)}\right)^2 \tag{2.9}$$

We can see variation with learning rate too, It will not have always same trend. As good learning rule, will have better cost function and so steeper loss throughout compared to in very high and low learning rate, it varies with number of iterations and so on with each iterations. The plot is figure 2.3 for learning rate with loss.

2.2 Discrete-Variable Artificial Neural Network

A discrete-variable network has unitary operators of hidden layers, acts on input sequentially with positive tanspose map,

The network output can be write as the composite function:

$$\rho^{out} = \xi^{out}\left(\xi^{N}\left(\cdots \xi^{2}\left(\xi^{1}\left(\rho^{in}\right)\right)\cdots\right)\right) \tag{2.10}$$

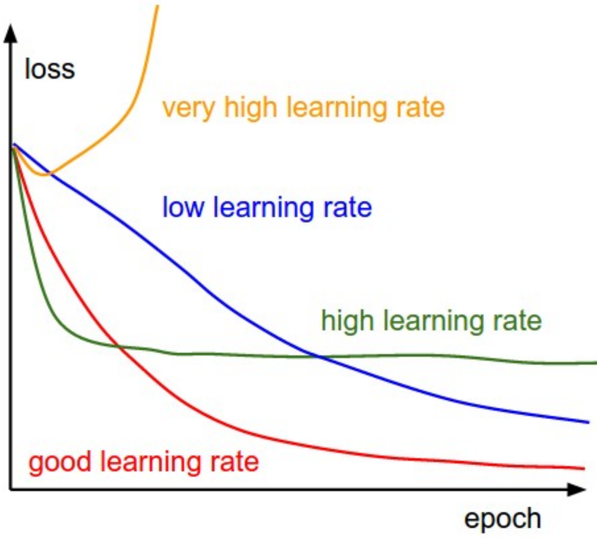

Figure 2.3: Loss vs Epoch for different learning rate

[Bose 19]

Cost function can be write by matching rate of learning network,

$$\begin{aligned} &\mathcal{M}\left(\rho, label\right) \\ =&\frac{1}{\left|\mathcal{D}_{train}\right|}\sum_{\mathcal{D}_{train}}\left(Pr\left(0_O\right|0_{EO}\right)+Pr\left(1_O\right|1_{EO}\right)) \\ =&1-\frac{1}{\left|\mathcal{D}_{train}\right|}\sum_{\mathcal{D}_{train}}\left(Pr\left(1_O\right|0_{EO}\right)+Pr\left(0_O\right|1_{EO}\right)) \end{aligned} \tag{2.11}$$

0_O tells the separable state, and 1_O denotes systems are entangled, and similarly for y_O(labels for the evaluated output), y_{EO}(for the expected output). From Matching rate we can also defined some parameters like loss function, the cost function of the assigned networks.
The loss function from the expected outputs,

$$L=\frac{1}{\mathcal{D}_{train}}\sum_{\mathcal{D}_{train}}\left[N'_O-N_{EO}\left(\rho_{|\varphi\rangle}\right)\right]^2 \tag{2.12}$$

where N'_O is the output of the neural network.

2.3 Continuous-Variable Artificial Neural Network

The continuous-variable system can be analyses by assuming the vectors of density matrix in their infinite dimensional Hilbert space. Qunatum Harmonic oscillator is one of the most common example of continuous-variable quantum system, and these solution of QHO in these states, known as Gaussian state, which has probability distribution.

So for 2 continuous-variable system, we can take analogue two-quantum mode Gaussian state, then to find entanglement between two continuous-variable system, we need to detect these two-qumode Gaussian states are separable or entangled. And the criteria is following for two states ($\hat{x}_1$,$\hat{p}_1$) and ($\hat{x}_2$, $\hat{p}_2$):

$$\left\langle (\Delta\hat{\lambda})^2 \right\rangle + \left\langle (\Delta\hat{\delta})^2 \right\rangle \geq c^2 + \frac{1}{c^2} + M^2, \tag{2.13}$$

$$\begin{aligned} \hat{\lambda} &= |c|\,\hat{x}_1 + \hat{x}_2/c \\ \hat{\delta} &= |c|\,\hat{p}_1 - \hat{p}_2/c \\ M &= |c|\sqrt{\left\langle (\Delta\hat{x}_1)^2 \right\rangle + \left\langle (\Delta\hat{p}_1)^2 \right\rangle - 1} \\ &\quad - \sqrt{\left\langle (\Delta\hat{x}_2)^2 \right\rangle + \left\langle (\Delta\hat{p}_2)^2 \right\rangle - 1}/|c|. \end{aligned} \tag{2.14}$$

Label sets to 0 value, and then check whether this inequality holds or not. If this inequality holds, the state is said separable and vice-versa.

Chapter 3

Artificial Neural Networks in Quantum Information

3.1 Entanglement for Qubits-Qubits system

The density matrix ρ for the Qubit(bipartite quantum systems) written as,

$$\rho = \sum_{i_1 i_2 j_1 j_2} \rho_{i_1 j_1, i_2 j_2} |i_1 j_1\rangle \langle i_2 j_2| \tag{3.1}$$

Applied PPT criterion as follows[PENG-HUI QIU 19],

$$\rho^{PT}_{i_1 j_1, i_2 j_2} = \rho_{i_1 j_2, i_2 j_1} \tag{3.2}$$

$$\rho^{R}_{i_1 j_1, i_2 j_2} = \rho_{i_1 i_2, j_1 j_2} \tag{3.3}$$

By matrix rearrangement for bipartite quantum state ρ to be separable if

$$sup\|T^2_{ij}\rho\|_1 \leqslant 1, \quad i, j \in \{1,..,4\} \tag{3.4}$$

T^n_{ij} is the transposition of elements in general tensor-product space $H^{\otimes n}$. For 2 spin- 1/2 states, density

$$\rho = p|\varphi_1\rangle\langle\varphi_1| + (1-p)|\varphi_2\rangle\langle\varphi_2| \tag{3.5}$$

where $|\varphi_1\rangle = a|00\rangle + b|11\rangle$, $|\varphi_2\rangle = a|10\rangle + b|01\rangle$. Plot the value of $f = \|T^2_{ij}\rho\|$ with respect to coefficient ”a”.

If $f > 1$, the state is said to be entangled, otherwise it is separable.

We made 3 neural networks to study the efficiency for the entangled and separable state. For that we used 7000 two qubit states [3500 entangled and 3500 separable]. Keep the Neural networks size as follow [25-25-5], [15-15-10-1], [10-10-5-5-2].

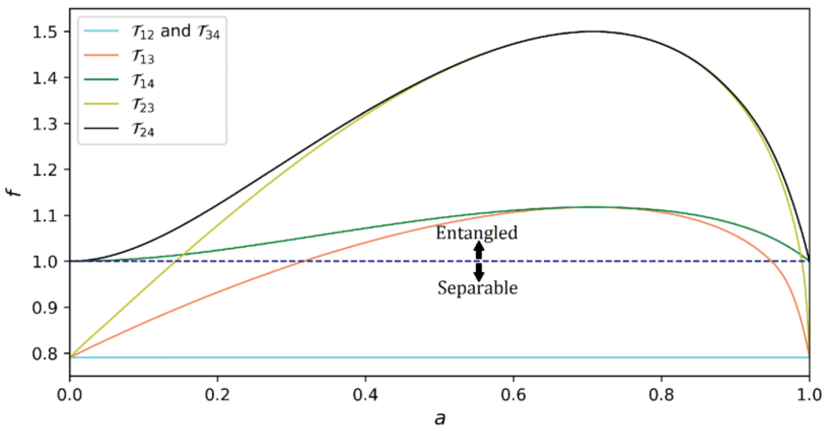

Figure 3.1: f vs a for entanglement and separability

[PENG-HUI QIU 19]

3.2 Entanglement for Qutrits

The partial positive transposition condition to determines entanglment for a general state of two qutrits fails, as it is being system whose states belong to 3-D Hilbert space. However, to provide the criterion for entanglement in the higher dimensional Hilbert-space, we have different aprroach.

Like a joint state of a composite system is said to be separable if it can be decomposed into a mixture of product states for their constituents. we consider the states of two qutrits which are a mixture of both individually the maximally mixed state and a maximally entangled state. These states are the generalization of Werner state for two qubits. Then the two-qutrit system is supposed to be separable if and only if the probability outcome for the maximally entangled state should not exceed $1/4$.

For a qutrit, if $|1\rangle$, $|2\rangle$and$|3\rangle$ are an orthonormal basis, then pure state $|\psi\rangle$ can be written as by normalization and overall phase,[Carlton M. Caves]

$$|\psi\rangle = e^{i\zeta_1} \sin\theta \cos\phi |1\rangle + e^{i\zeta_2} \sin\theta \sin\phi |2\rangle + \cos\theta |3\rangle \tag{3.1}$$

$$\begin{aligned} 0 \le \theta, \phi \le \pi/2 \\ 0 \le \zeta_1, \zeta_2 \le 2\pi \end{aligned} \tag{3.2}$$

We can look the matrix representations of hermitian operators of qutrits state, which is SU(3) group generator:[Nielsen 00]

$$U = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}$$

$$u_1 = \begin{pmatrix} 1 & 0 & 0 \\ 0 & \omega & 0 \\ 0 & 0 & \omega^2 \end{pmatrix}, u_2 = \begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 1 & 0 & 0 \end{pmatrix}$$

$$u_3 = \begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & \omega \\ \omega^2 & 0 & 0 \end{pmatrix}, u_4 = \begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & \omega^2 \\ \omega & 0 & 0 \end{pmatrix}$$

$$u_5 = u_1^\dagger, u_6 = u_2^\dagger$$

$$u_7 = u_3^\dagger, u_8 = u_4^\dagger$$

This representation of qutrits state is complete and a state can be written as,

$$\rho = -1 + \sum_{m=1}^{4} \sum_{k=0}^{2} p(m,k)|mk\rangle\langle mk| \tag{3.3}$$

For maximally mixed state(M_9) and maximally entangled state $|\Psi\rangle$

$$M_9 = \frac{1}{9} I \otimes I \tag{3.4}$$

$$|\Psi\rangle = \frac{1}{\sqrt{3}}(|1\rangle \otimes |1\rangle + |2\rangle \otimes |2\rangle + |3\rangle \otimes |3\rangle) \tag{3.5}$$

These mixture will have form,

$$\rho_\epsilon = (1-\epsilon)M_9 + \epsilon|\Psi\rangle\langle\Psi| \tag{3.6}$$

$0 \leq \epsilon \leq 1$. These states of the two qutrits will be separable if it can be expressed as an ensemble of the product states. So we can show it that the state is separable, if and only if $\epsilon \leq 1/4$.

We can look at expansion for maximally entangled state, [Carlton M. Caves]

$$|\Psi\rangle\langle\Psi| = \sum_{a,b} |a\rangle\langle b| \otimes |a\rangle\langle b|$$

$$= \frac{1}{9}\Big(I \otimes I + \frac{3}{2}\,(u_1 \otimes u_1 - u_2 \otimes u_2 + u_3 \otimes u_3$$

$$+u_4 \otimes u_4 - u_5 \otimes u_5 + u_6 \otimes u_6 - u_7 \otimes u_7 + u_8 \otimes u_8) \tag{3.7}$$

For mixed state;

$$\rho_\epsilon = \frac{1}{9}\left(I \otimes I + \frac{3\epsilon}{2}\left(u_1 \otimes u_1 - u_2 \otimes u_2 + u_3 \otimes u_3\right.\right.$$

$$\left.\left.+u_4 \otimes u_4 - u_5 \otimes u_5 + u_6 \otimes u_6 - u_7 \otimes u_7 + u_8 \otimes u_8\right)\right) \tag{3.8}$$

The expansion coefficient from equation 3.3 of density matrix of qutrits,

$$c_{\alpha\beta} = \frac{9}{4}\operatorname{tr}\left(\rho u_\alpha \otimes u_\beta\right) \tag{3.9}$$

Density matrix of qutrits,

$$\rho = \frac{1}{9} c_{\alpha\beta} u_\alpha \otimes u_\beta \tag{3.10}$$

Coefficients would be,

$$c_{0j} = c_{j0} = 0, \quad c_{jk} = 0 \quad \text{for } j \neq k$$

$$c_{11} = -c_{22} = c_{33} = c_{44} = -c_{55} = c_{66} = -c_{77} = c_{88} = \frac{3\epsilon}{2} \tag{3.11}$$

Chapter 4

Artificial Neural Networks in Nuclear Magnetic Resonance

Artificial neural network(ANN) analyses the NMR spectroscopy with different approach which deals data quantification. It does lineshape fitting analysis, and performance has been much better through ann models. When we talk about quantification of lipoprotein in different density like(very low, intermediate, low, high density), then ANN increases the value of ^1H NMR lipoprotein quantification to the extends where it could be one's choice in advanced reserach setting.

4.1 Lineshape fitting(LF) analysis

Lineshape fitting is one of the reliable quantification of overlapping resonances got from spectra of 'H NMR. So the model for the different density of lipoprotein can be write in form of characteristic parameters which is individual Lorentzians Li(can be line widths, chemical shifts and intensities);

$$\begin{aligned} M_{\mathrm{VLDL}}(f) &= \mathrm{L}_1(f)^{\mathrm{VLDL}} + \mathrm{L}_2(f)^{\mathrm{VLD}} + \mathrm{L}_3(f)^{\mathrm{VLD}} \\ \mathrm{M}_{\mathrm{IDL}}(f) &= \mathrm{L}_1(f)^{\mathrm{IDL}} + \mathrm{L}_2(f)^{\mathrm{IDL}} + \mathrm{L}_3(f)^{\mathrm{ID}} \\ \mathrm{M}_{\mathrm{LDL}}(f) &= \mathrm{L}_1(f)^{\mathrm{LDL}} + \mathrm{L}_2(f)^{\mathrm{LDL}} + \mathrm{L}_3(f)^{\mathrm{LDL}} \\ \mathrm{M}_{\mathrm{HDL}}(f) &= \mathrm{L}_1(f)^{\mathrm{HDL}} \end{aligned} \tag{4.1}$$

where $M_{VIDL}(f)$ is Model for very low density lipoprotein, and so on for intermediate, low, and high density in the above expressions.

Then the model of the methyl resonance from a plasma spectrum can be look as,

$$\begin{aligned} \mathrm{M_{PLASMA}}(f) =& \mathrm{M_{VLDL}}(f) + \mathrm{M_{IDL}}(f) + \mathrm{M_{LDL}}(f) + \mathrm{M_{HDL}}(f) \\ & + \mathrm{c}_0 + \mathrm{c}_1 f \end{aligned} \tag{4.2}$$

$c_0 + c_1 v$ is the background resonance due to residual water, or fatty acids or proteins.

4.2 Lipoprotein isolation by ANN

Here we can use sigmoid or linear function as the activation function

$$f\left(Z^i\right) = \begin{cases} \text{linear}(Z^i) & i \neq N_f \\ \text{sigmoid}\left(Z^i\right) = \frac{1}{1+e^{-Z^i}} & i = N_f \end{cases}$$

Each neuron is first multiplied by their corresponding weighting factor which influence the coming output, From different study group of 59 cases: 37 spectra choose to for the training set and rest 22 spectra for the test data set.[M. Ala-Korpela 95]

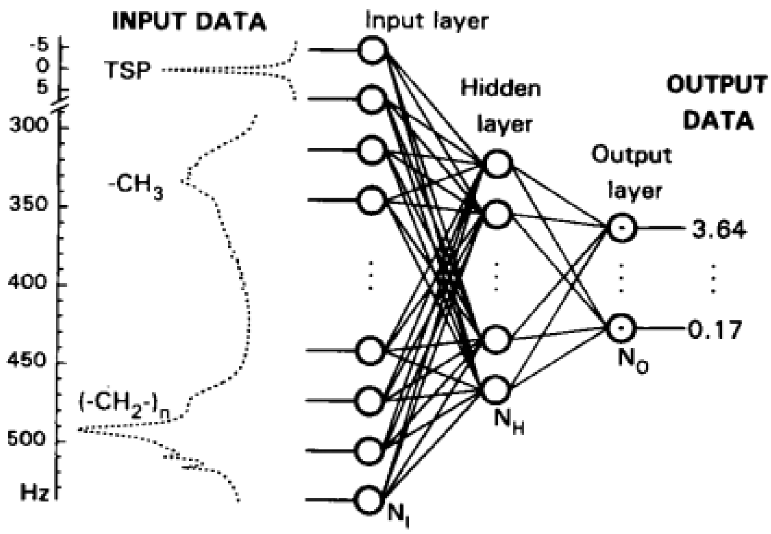

Figure 4.1: spectra of lipoprotein and analysed with ann networks
[M. Ala-Korpela 95]

The number for input neurons (Nl) was kept around 950, hidden neurons (NH) was kept between 0 to 100, and outputs(No) was kept 20. For the different no. of hidden layers it has observed isolation of lipoprotein in very good extends.

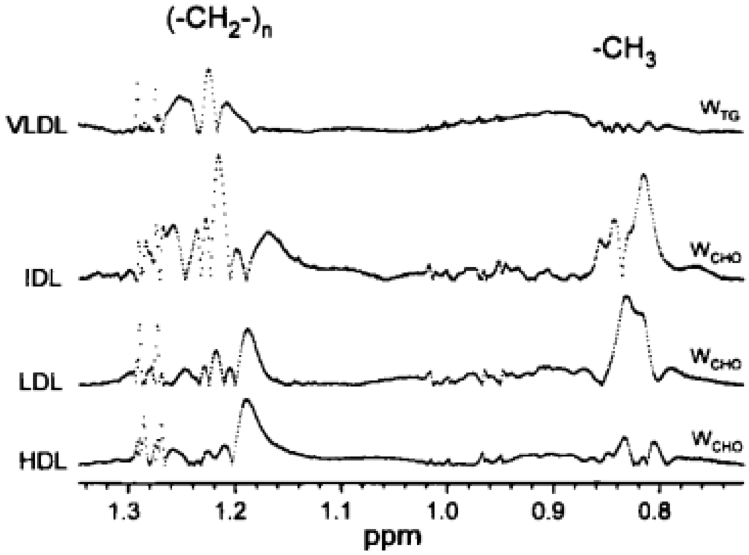

Figure 4.2: lipoprotein of different density

[M. Ala-Korpela 95]

With 960 inputs, 20 outputs and 1000 iterations have made separation or isolation between density of protein.

Chapter 5

Analysis & Results

5.1 Efficiency for Bipartitie Quantum Systems

For the following 3 neural networks [25-25-5], [15-15-10-1], [10-10-5-5-2] to study the efficiency for the entangled and separable state. And the given 7000 two qubit states [3500 entangled and 3500 separable]. Parameters: Activation(sigmoid) function used, $\mathrm{sigmoid}\left(Z^i\right) = \frac{1}{1+e^{-Z^i}}$, Gradient descent otimization for loss and backward propagation.

	Network 1		Network 2		Network 3	
	Entangled	Separable	Entangled	Separable	Entangled	Separable
Efficiency	50.634	50.556	57.978	42.72	49.98	50.927

Rest of efficiency table for Qubit has been in appendix table-1.

[15-15-10-1] network stands for a network with 15 layers (input-15 hidden-output) with 15, 15, 10 and 1 neurons in the four hidden layers from input to output order. So more no. of hidden layers make good enough efficient model upto an extent, but also more input biased and weighted function done their job.

The cost function is one half of the mean squared error

$$J\left(\theta_0, \theta_1\right) = \frac{1}{2N} \sum_{i=1}^{N} \left(f_\theta\left(x^{(i)}\right) - y^{(i)}\right)^2 \tag{5.1}$$

And loss function is defined as,

$$L = \frac{1}{\mathcal{D}_{train}} \sum_{\mathcal{D}_{train}} \left[N'_{AR} - N_{AR}\left(\rho_{|\varphi\rangle}\right)\right]^2 \tag{11}$$

Loss function for Qubit-Qubit entanglement model converges after around 600 iterations, and test loss is more compare to train loss for at any time of iterations.

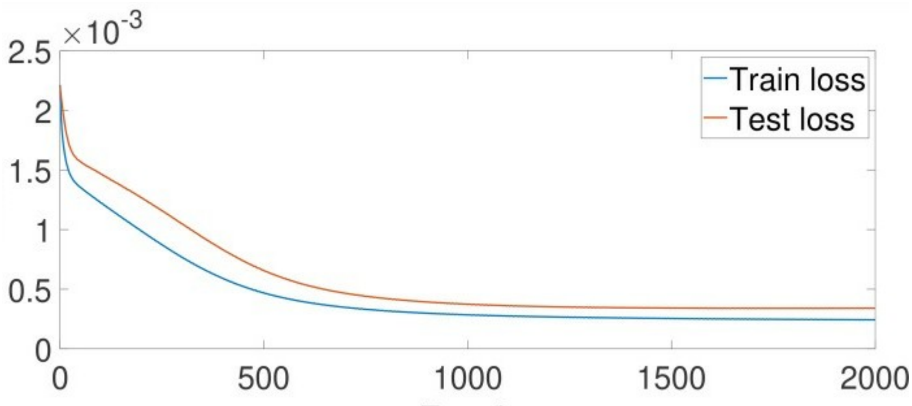

Figure 5.1: Loss Function vs Iteration

we looked the loss function for continuous-variable ann over 2000 epoch for training and testing data sets,

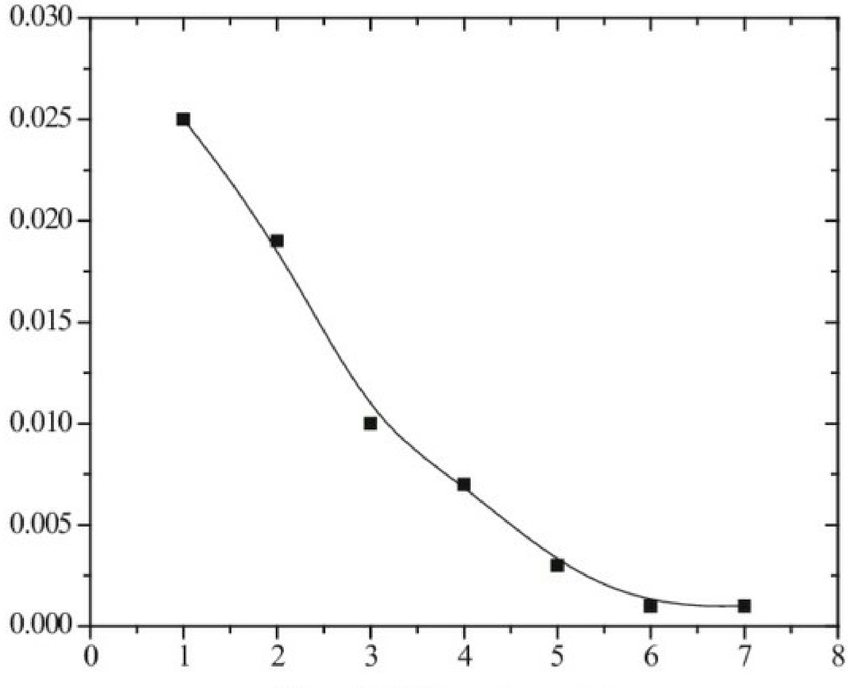

Figure 5.2: Mean Square error vs No. of hidden layers of neuron

Convergence of the training loss for the ann model with 6 hidden layers. As by increasing more hidden layers we can have more efficient model, but computational time would be around double by increasing one more hidden layers. So we can fix our hidden layers till we get a convergence of mean square error.

5.2 Efficiency for Qutrits

For the following 3 neural networks [25-25-20-5], [20-18-16-5], [16-12-12-5-2] to study the efficiency for the entangled and separable state. And the given 1500 two qutrit states [750 entangled and 750 separable]. We have here only those state of qutrit which are mixed state of maximally entangled state and maximally mixed state.

[25-25-20-5] network stands for a network with 25 layers (input-25 hidden-output) with 25, 25, 20 and 5 neurons in the four hidden layers from input to output order. Network has more than 90% efficiency, Network-2 has around 80% and Network-3 has around 75%. We goes for more number of inputs state rather than more hidden layers for qutrits state.

	Network 1		Network 2		Network 3	
	Entangled	Separable	Entangled	Separable	Entangled	Separable
Efficiency	98.6	97.82	82.47	80.76	74.62	76.93

Rest 25 efficiency for different epochs of Qutrit-Qutrit entanglement has mentioned in Table-2 of appendix. We see Network 1 is higher efficiency compare to Network-2, and 3. But in Qubit [15-15-10-1] hidden layer network has much better efficiency.

Validation or test data has been less loss means better trained for qutrit-qutrit entanglement and got an convergence.

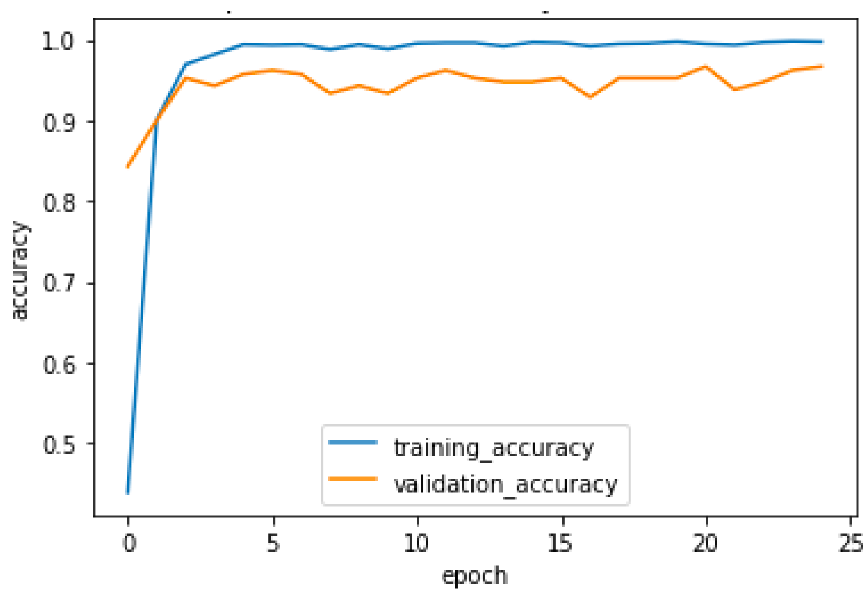

Figure 5.3: Accuracy vs Epochs for [25-25-20-5] networks

Training and Validation accuracy graph for [25-25-20-5] networks can been seen against number of epochs.

Chapter 6

Conclusion and Future work

- We have seen good correlation between number of epochs and hidden layers of networks, as below from 200 epochs faster drops, and for higher number of layers drop becomes slow. So need to fit an optimal solution in consideration of those factors.
- Continuous-variable has 92% detecting rate of entanglement, which has very much better than discrete-variable networks for 6 hidden layers in Qubit-Qubit systems.
- In Qutrit-Qutrit entanglement [25-25-20-5] network has good efficiency compare to network having more number of hidden layers, as qutrit has more number of input variables, it's need more biased and weighted function compared to qubit. So in qubit system [15-15-10-1] has better efficiency, which is having more hidden layers.
- Isolation of lipoprotein has been observed for Input variables around 900 and output variables 20, where spectra frequency belonged between 300-500Hz.
- ANN model best can work on quantification of NMR spectra with analyses of phase shift, peak and identification of various peak got from spectra.

Chapter 7

Appendix

7.1 Data 1- Efficiency for qubit-qubit entanglement

Epochs	Network 1		Network 2		Network 3	
	Entangled	Separable	Entangled	Separable	Entangled	Separable
	50.634	50.566	57.978	42.72	49.98	50.927
	49.75	49.42	61.26	39.57	50.142	50.46
	56.919	44.787	58.642	42.17	50.55	50.678
	52.734	47.16	50.448	49.425	54.214	47.028
	45.9	52.42	56.882	44.37	51.784	49.45
	51.162	48.168	63.66	36.47	48.88	52.07
	56.56	46.38	64.487	35.79	50.47	48.16
	52.482	46.446	57.62	42.44	61.591	40.124
	48.918	50.47	50.78	51.21	48.574	51.79
	50.725	50.573	47.48	53.37	50.57	51.46
	52.27	47.49	62.428	38.42	40.36	58.452
	54.462	48.28	66.67	33.94	52.268	47.38
	53.54	45.97	56.345	43.216	48.076	53.132
	46.79	53.02	58.691	41.426	48.87	52.238
	51.93	50.78	40.38	59.952	55.79	45.147

Neural networks size as follow Network1-[25-25-5], Network2-[15-15-10-1], Network3-[10-10-5-5-2].

7.2 Data 2- Efficiency for qutrit-qurit entanglement

Epochs	Network 1		Network 2		Network 3	
	Entangled	Separable	Entangled	Separable	Entangled	Separable
	98.6	97.82	82.47	80.76	74.62	76.93
	95.88	98.3	79.67	81.74	66.42	60.46
	98.32	96.81	83.67	80.71	69.45	72.86
	94.734	97.16	83.58	81.64	73.24	71.04
	100	96.42	76.89	74.72	61.43	59.72
	96.14	95.39	76.67	76.43	64.86	62.72
	96.26	94.7	74.75	75.79	61.72	60.19
	97.93	96.46	72.06	70.42	61.34	62.45
	98.968	93.87	70.78	71.25	68.74	61.74
	96.785	96.579	77.46	79.7	70.73	65.45
	93.7	98.93	76.58	74.46	68.64	62.26
	91.4	98.45	79.75	73.46	63.58	62.65
	93.74	96.27	76.455	73.268	58.67	56.192
	93.92	95.62	78.64	71.44	68.43	62.86
	96.43	93.98	79.68	76.45	69.73	64.31
	97.02	94.49	75.17	75.32	65.16	61.72
	97.26	94.8	74.745	76.79	61.22	60.14
	96.92	95.66	71.86	70.76	61.24	62.41
	96.67	93.17	70.75	74.2	68.345	60.78
	97.235	95.169	77.76	79.27	70.733	64.45
	98.71	92.93	74.58	75.44	69.63	62.65
	90.42	96.45	73.75	73.467	65.52	64.68
	96.94	95.27	77.45	72.458	58.27	56.49
	94.05	98.42	72.64	70.44	65.43	62.85
	95.43	92.98	76.68	76.48	69.453	64.305

Neural networks size as follow Network1-[25-25-20-5], Network2- [20-18-16-5], Network3-[16-12-12-5-2]

7.3 Tripartite qutrits SLOCC representation

Name	Representative	Name	Representative
$\|\psi_0\rangle$	$\|000\rangle$	$\|\psi_{12}\rangle$	$\|000\rangle+\|011\rangle+\|101\rangle+\|112\rangle$
$\|\psi_1\rangle$	$\|000\rangle+\|011\rangle$	$\|\psi_{13}\rangle$	$\|000\rangle+\|011\rangle+\|112\rangle+\|120\rangle$
$\|\psi_2\rangle$	$\|000\rangle+\|011\rangle+\|022\rangle$	$\|\psi_{14}\rangle$	$\|000\rangle+\|011\rangle+\|120\rangle+\|101\rangle$
$\|\psi_3\rangle$	$\|000\rangle+\|101\rangle$	$\|\psi_{15}\rangle$	$\|000\rangle+\|011\rangle+\|120\rangle+\|102\rangle$
$\|\psi_4\rangle$	$\|000\rangle+\|110\rangle$	$\|\psi_{16}\rangle$	$\|000\rangle+\|011\rangle+\|022\rangle+\|101\rangle$
$\|\psi_5\rangle$	$\|000\rangle+\|111\rangle$	$\|\psi_{17}\rangle$	$\|000\rangle+\|011\rangle+\|022\rangle+\|101\rangle+\|112\rangle$
$\|\psi_6\rangle$	$\|000\rangle+\|011\rangle+\|101\rangle$	$\|\psi_{18}\rangle$	$\|000\rangle+\|011\rangle+\|022\rangle+\|112\rangle+\|120\rangle$
$\|\psi_7\rangle$	$\|000\rangle+\|011\rangle+\|112\rangle$	$\|\psi_{19}\rangle$	$\|000\rangle+\|011\rangle+\|022\rangle+\|120\rangle+\|101\rangle$
$\|\psi_8\rangle$	$\|000\rangle+\|011\rangle+\|120\rangle$	$\|\psi_{20}\rangle$	$\|000\rangle+\|011\rangle+\|122\rangle$
$\|\psi_9\rangle$	$\|000\rangle+\|101\rangle+\|202\rangle$	$\|\psi_{21}\rangle$	$\|000\rangle+\|110\rangle+\|220\rangle$
$\|\psi_{10}\rangle$	$\|000\rangle+\|111\rangle+\|202\rangle$	$\|\psi_{22}\rangle$	$\|000\rangle+\|111\rangle+\|220\rangle$
$\|\psi_{11}\rangle$	$\|000\rangle+\|111\rangle+\|201\rangle$	$\|\mathcal{G}\rangle$	$\|000\rangle+\|111\rangle+\|222\rangle$

Name	Representative
$\|\pi(\phi,\varphi,\chi,\psi)\rangle$	$\|000\rangle+\|011\rangle+\|1\phi\varphi\rangle+\|2\chi\psi\rangle$
$\|\phi_0\rangle$	$\|000\rangle+\|011\rangle+\|022\rangle+\|101\rangle+\|202\rangle$
$\|\phi_1\rangle$	$\|000\rangle+\|011\rangle+\|022\rangle+\|110\rangle+\|220\rangle$
$\|\varphi_1\rangle$	$\|000\rangle+\|011\rangle+\|022\rangle+\|101\rangle+\|212\rangle$
$\|\phi_2(\phi,\varphi)\rangle$	$\|000\rangle+\|011\rangle+\|101\rangle+\|112\rangle+\|2\phi\varphi\rangle$
$\|\varphi_2(\phi,\varphi)\rangle$	$\|000\rangle+\|011\rangle+\|112\rangle+\|120\rangle+\|2\phi\varphi\rangle$
$\|\phi_3(\phi,\varphi)\rangle$	$\|000\rangle+\|011\rangle+\|120\rangle+\|101\rangle+\|2\phi\varphi\rangle$
$\|\phi_4\rangle$	$\|000\rangle+\|011\rangle+\|101\rangle+\|112\rangle+\|202\rangle+\|221\rangle$
$\|\psi_{23}\rangle$	$\|000\rangle+\|011\rangle+\|101\rangle+\|112\rangle+\|210\rangle+\|202\rangle$
$\|\phi_5\rangle$	$\|000\rangle+\|011\rangle+\|101\rangle+\|112\rangle+\|221\rangle+\|210\rangle$
$\|s_0\rangle$	$\|000\rangle+\|011\rangle+\|112\rangle+\|120\rangle+\|202\rangle+\|221\rangle$
$\|\phi_6\rangle$	$\|000\rangle+\|011\rangle+\|112\rangle+\|120\rangle+\|221\rangle+\|210\rangle$
$\|\psi_{24}\rangle$	$\|000\rangle+\|011\rangle+\|120\rangle+\|101\rangle+\|221\rangle+\|210\rangle$
$\|\phi_7\rangle$	$\|000\rangle+\|011\rangle+\|022\rangle+\|101\rangle+\|112\rangle+\|202\rangle+\|221\rangle$
$\|\phi_8\rangle$	$\|000\rangle+\|011\rangle+\|022\rangle+\|101\rangle+\|112\rangle+\|210\rangle+\|202\rangle$
$\|s_1\rangle$	$\|000\rangle+\|011\rangle+\|022\rangle+\|101\rangle+\|112\rangle+\|221\rangle+\|210\rangle$
$\|w_0\rangle$	$\|000\rangle+\|011\rangle+\|022\rangle+\|101\rangle+\|112\rangle+\|202\rangle$
$\|\varphi_3\rangle$	$\|000\rangle+\|011\rangle+\|022\rangle+\|101\rangle+\|112\rangle+\|220\rangle$
$\|\phi_9\rangle$	$\|000\rangle+\|011\rangle+\|022\rangle+\|101\rangle+\|112\rangle+\|221\rangle$

Figure 7.1: Tripartite qutrits state representation

[Honda 01]

Bibliography

[Bose 19] George V. Bose. *Useful Plots to Diagnose your Neural Network*. 2019.

[Carlton M. Caves] Gerard J. Milburn Carlton M. Caves. *Qutrit Entanglement*.

[Honda 01] Kentaro Honda. *Graphical Classification of Entangled Qutrits*. 201.

[M. Ala-Korpela 95] J. D. Bell M. Ala-Korpela Y. HiltunenS. *Lipoprotein lipid profiles from1H NMR data of human plasma*. 1995.

[Nielsen 00] Michael A. Nielsen & Isaac L. Chuang. Quantum computation and quantum information. 2000.

[PENG-HUI QIU 19] YI-WEI SHI PENG-HUI QIU XIAO-GUANG CHEN. *Detecting Entanglement With Deep Quantum Neural Networks*. July 2019.

[Wittek 14] Peter Wittek. *Quantum Machine Learning*. 2014.