Triangle Network Nonlocality using Neural Network Oracle

A Dissertation Submitted

in Partial Fulfilment of the Requirements

for the Degree of

MASTER OF SCIENCE

in

PHYSICS

by

ANANTHA KRISHNAN S

(Roll No. IMS18022)

Under the Guidance of

Prof. Anil Shaji

School of Physics

IISER Thiruvananthapuram

to

SCHOOL OF PHYSICS

INDIAN INSTITUTE OF SCIENCE EDUCATION AND RESEARCH

THIRUVANANTHAPURAM - 695 551, INDIA

January 2024

DECLARATION

I, **Anantha Krishnan S (Roll No: IMS18022)**, hereby declare that, this report entitled **"Triangle Network Nonlocality using Neural Network Oracle"** submitted to Indian Institute of Science Education and Research Thiruvananthapuram towards partial requirement of **Master of Science** in **Physics**, is an original work carried out by me under the supervision of **Prof. Anil Shaji** and has not formed the basis for the award of any degree or diploma, in this or any other institution or university. I have sincerely tried to uphold the academic ethics and honesty. Whenever an external information or statement or result is used then, that have been duly acknowledged and cited.

Thiruvananthapuram - 695 551 **Anantha Krishnan S**

April 2023

CERTIFICATE

This is to certify that the work contained in this project report entitled **"Triangle Network Nonlocality using Neural Network Oracle"** submitted by **Anantha Krishnan S** (**Roll No: IMS18022**) to Indian Institute of Science Education and Research, Thiruvananthapuram towards the partial requirement of **Master of Science** in **Physics** has been carried out by him under my supervision and that it has not been submitted elsewhere for the award of any degree.

Thiruvananthapuram - 695 551 Prof. Anil Shaji

April 2023 Project Supervisor

ABSTRACT

Witnessing nonlocal correlations lies at the core of the foundational aspects of quantum mechanics, and such correlations have huge applications in information processing. In the case of simple bipartite quantum systems, we can achieve this by using Bell inequalities. But as soon as the scenario grows in complexity, relying on the Bell inequalities becomes unfeasible. This is where machine learning can provide an alternative for detecting and quantifying non-locality. Since machine learning structures such as neural networks are classical, it is possible to show if a quantum distribution is nonlocal or not by letting the machine learn it and make it reproduce the distribution, i.e. we are using classical resources to reproduce an observed probability distribution that arises from measurements on a quantum state. If the distribution is local, the machine can learn it; if nonlocal the machine can't reproduce it, and we can try to use this disparity to create a measure for nonlocality. We are focusing on the triangle case of nonlocality which was proposed by Renou[1] which shows a novel kind of nonlocality that does not require measurement inputs, unlike the standard Bell nonlocality.

The main objectives are (a) Using this approach to classify states as nonlocal or not and further find non-maximally entangled states which are nonlocal in the triangle setting. (b) Expanding on the work by Krivachy[2] and Renou[1] to X-quantum states that is a bigger set than the already studied Bell states. (c) Improving the optimization, as optimization plays a huge role since its a huge sample space. (d) Finding the best measurement setting for X states and eventually a measure for nonlocality in the triangle setting. (c) Expanding on the noise robustness study by Krivachy[2] as a possible means to map the nonlocal boundary in the triangle network case.

Contents

List of Figures

1 Introduction

The project will first introduce the concept of nonlocality for the bipartite case that can be witnessed using the bell theorem, then move on to the more complex triangle network case proposed by Renou et al[1]. Then we touch upon why it is important to use machine learning for distinguishing nonlocal states from local ones and get started on how we can attempt to solve this problem using neural networks. We generalize this method which has been used on Bell states to X states and we found the best measurement setting for particular sets of X states which turns out to be entangled measurements. We further consider the possibility of the consistent behavior shown by comparing Euclidean distances and measurement parameters of different nonlocal distributions in the X states with entangled measurements as proof of some property in the 64-dimensional Hilbert space.

1.1 Standard Bell Nonlocality

"Bell's theorem is the most profound discovery of science" – Henry Stapp. This is because it showed that the prediction of quantum theory is inconsistent with physical theories with a natural notion of locality; thus undermining our centuries-old understanding of physical reality.

We can understand this through an experiment[3] consisting of two systems that may have previously interacted(maybe produced by a common source) and then spatially separated(even space-like) by two agents Alice and Bob for measurement. Before getting separated, a referee gives Alice and Bob a random bit x, y each of which decides the measurement out of Alice and Bob's different measurement options. After the measurement, the outcomes are a and b. And on repeating this, we get a probability distribution $p(ab|xy)$.

Doing this practically shows that

$$p(ab|xy) \neq p(a|x)p(b|y) \tag{1}$$

this means correlations are present, which may not necessarily be a direct influence of one on the other but could be due to some dependence relation between the 2 systems that were established when they interacted in the past.

So if we include all the past factors the residual indeterminacies about the output should be decoupled[4].

$$p(ab|xy) = \int_{\lambda} d\lambda \; q(\lambda)p(a|x,\lambda)p(b|y,\lambda) \tag{2}$$

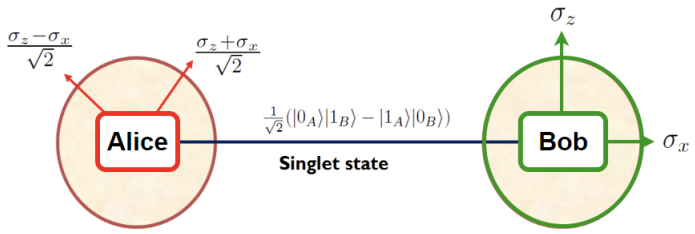


Figure 1: Alice's and Bob's measurement choice

All classical systems being local theories will obey the above equation, as long as we consider all hidden variables, if any. Assuming that Quantum mechanics is also a local theory it should do the same. But in 1964 Bell came with an astounding theorem that undermined this assumption. The thought experiment he devised gave us an inequality, which when violated implies nonlocal behavior, and it has been found that quantum states, for certain measurements, violate this inequality and show nonlocal behavior.

In the experiment, we set the measurement choices $x, y \in 0, 1$ and the measurement outcomes $a, b \in -1, 1$ respectively, and Alice and Bob cannot communicate with each other.

We can then define an expression

$$\begin{aligned} |S| &= |a_0b_0 + a_1b_0 + a_0b_1 - a_1b_1| \\ |\langle S\rangle| &= |\langle a_0b_0\rangle + \langle a_0b_1\rangle + \langle a_1b_0\rangle + \langle a_1b_1\rangle| \leq 2 \ , \text{ where } \langle a_xb_y\rangle = \textstyle\sum_{a,b} ab\ p(ab|xy) \end{aligned} \tag{3}$$

which will be ≤ 2 if it satisfies (2)

This is the CHSH version of the Bell's No-go theorem, for any theory satisfying the condition of local realism we have $|\langle S\rangle| \leq 2$.

Quantum violation of Bell's inequality:

Here Alice and Bob can share some bipartite state

$\rho_{AB} = |\psi^-\rangle_{AB}\langle\psi^-| \in D(C^2 \otimes C^2);\ |\psi^-\rangle_{AB} := \frac{1}{\sqrt{2}}(|01\rangle_{AB} - |10\rangle_{AB})$

where the CHSH quantity is

$$|\langle S\rangle| = |\langle a_0b_0\rangle + \langle a_0b_1\rangle + \langle a_1b_0\rangle + \langle a_1b_1\rangle|$$

where now $\langle a_xb_y\rangle = -\vec{x}.\vec{y}$

and for this measurement setting in this bipartite state, we get $|\langle S\rangle| = 2\sqrt{2} > 2$ which means the correlations present in the quantum state are nonlocal in nature.

1.2 Quantum Nonlocality in the Triangle network

Quantum networks have been shown to result in completely novel forms of quantum correlations [1]; i.e quantum nonlocality can be demonstrated using only the joint statistics of fixed local measurement outputs without the need of various input settings like in the case of the standard Bell nonlocality. That is we dont require any of the random measurement choices to ensure nonlocal correlation.

Here we consider the triangle quantum network which consists of three agents or observers(Alice, Bob and Charlie) and every pair of agents is connected by a bipartite qubit source; thus providing a shared physical system. Importantly, the three observers share no common (i.e. tripartite) piece of information, as the three sources are assumed to be independent of each other. Each observer provides an output based on the received physical resources (a, b and c, respectively). Contrary to standard Bell nonlocality tests, the observers receive no input for which measurement setting to use. The statistics of the experiment are thus given by the joint probability distribution P(a, b, c).

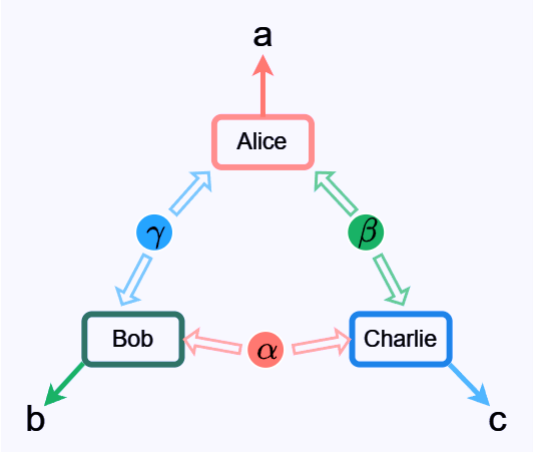


Figure 2: The Triangle network[1]

The family of quantum distributions $P_Q(a, b, c)$ is constructed using both entangled quantum states and entangled joint measurements. Take note that entangled measurements are different from the kind of measurements that we used in the Standard Bell Nonlocality. The entangled measurements done by agents A, B or C acts on the two shared qubits of two separate bipartite qubit sources. The sources giving entangled qubit states as well as the act of measuring two bipartite sources by one of the agents using entanglement measurements is responsible for the nonlocal behavior. These distributions are nonlocal if they cannot be represented by

$$P_Q(a, b, c) = \int d\alpha \int d\beta \int d\gamma P_A(a|\beta, \gamma) P_B(b|\gamma, \alpha) P_C(c|\alpha, \beta) \tag{4}$$

where $\alpha \in X, \beta \in Y$ and $\gamma \in Z$ are three local variables distributed by each source.

$P_A(a|\beta, \gamma), P_B(b|\gamma, \alpha)$ and $P_C(c|\alpha, \beta)$ are the response functions for Alice, Bob, and Charlie.

Here we don't have a convenient method such as Bell inequalities to classify states as nonlocal or local, since in complex and higher dimensional cases like these the boundary between nonlocal and local spaces are non-convex, it is difficult to find something like the Bell inequality. Analytical proof for the existence of nonlocal behavior already exists, Renou [1] had proven that there are indeed states

that exhibit nonlocality by using logical contradiction. The necessary properties that any trilocal model should have to reproduce $P_Q(a, b, c)$ were found unable to be satisfied simultaneously. This shows the distribution has some intrinsic randomness since it showed nonlocal behavior without relying on random measurement choices. The disadvantage of this distribution is that it assumes all the sources as the same and independent; experimentally keeping the three sources independent can be demanding.

These examples fundamentally differ from the standard bell case and its triangle adaptation done by Fritz [5].

1.3 Machine learning

For bipartite cases, we can classify nonlocality using Bell inequalities, but when its systems grow in complexity like in n-qubit systems and triangle networks we can no longer rely on Bell-type inequalities. But such casual inference task is challenging for both analytical and standard numerical techniques. In these scenarios, it has been shown[2] that using machine learning as a numerical tool to learn the classical strategies required to reproduce a distribution and using the trained model as a classifier works. As such, a neural network acts as the oracle for the observed behavior, demonstrating that it is classical if it can be learned through training the neural network.

2 Methodology

Here we are dealing with decentralized causal structures, where several independent sources are shared among the parties over a network. In such complex networks, the boundary between local and nonlocal correlations becomes nonlinear, and the local set is non-convex. Here like the work by Renou[1] we can encode the casual structure into a neural network and train the network to reproduce the target distribution; this shows if the "local casual model is learnable". And so if the Neural network can learn and reproduce the target distribution it is local.

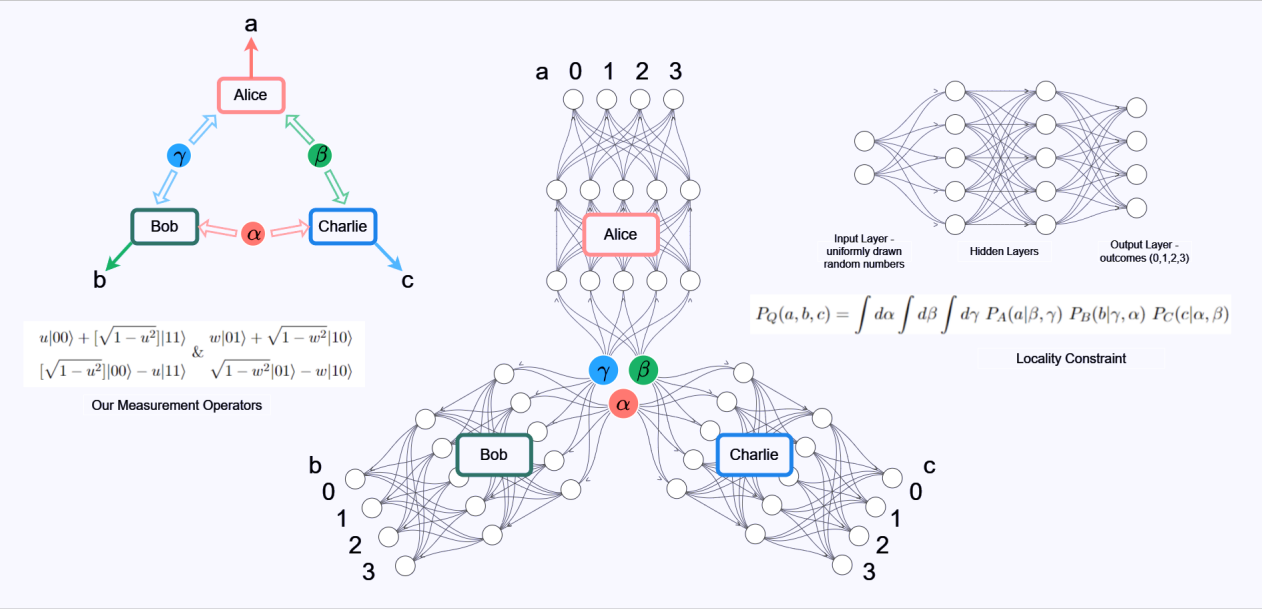


Figure 3: Triangle network and its neural network encoding

If the target distribution is outside the local set then the machine cannot reproduce it, this is because we haven't included a definite structure to recreate nonlocal correlations i.e. the neural network model is a classical structure incapable of reproducing quantum properties. So in the case of nonlocal target distributions, the Neural network approximates to the closest local distribution as per the local constraint(4).

To have a quantitative study on this we incorporate noise into the distribution. Adding sufficient noise to a nonlocal distribution brings it to the local set. So the methodology used by [2], was to start from zero visibility(max noise parameter) of the target distribution and gradually increase the visibility and as it leaves the local set we get an idea of the value of the visibility when it touches the boundary of local and nonlocal spaces. Adding noise also gives us a control over the behavior of nonlocality in this

situation.

The structure consists of three sources α, β, and γ which send information through either a classical or quantum channel to three parties A B and C.

We now convert the experiment that had three bipartite qubit sources α, β, and γ to three numerical sources that act as inputs for our machine model. Since we had three independent sources we are taking three independent values from a uniform distribution between 0 and 1.

Each party's response function is represented by a constraint-connected multilayer perceptron with rectified linear or tangent hyperbolic activations and softmax layer at the last layer to ensure the probabilities are normalized. As we said, inputs to the three perceptrons - hidden variables n_1 for α, n_2 for β, and n_3 for γ are random variables drawn from a uniform distribution on the continuous interval between 0 and 1.

No	n_1	n_2	n_3	$p(a\|n_2, n_3)$	$p(b\|n_1, n_3)$	$p(c\|n_1, n_2)$	p(a,b,c)
1	0.9679	0.9684	0.4667	...	...	...	...
2	0.6387	0.0864	0.7723	...	...	...	...
3	0.3483	0.0116	0.7508	...	...	...	...
:	:	:	:	:	:	:	:
1000	0.0575	0.1282	0.9497	...	...	...	...

Due to the communication constraints of the triangle network, the inputs are routed to the perceptron in a restricted manner. The network constraint being

$$P_Q(a,b,c) = \int dn_1 \int dn_2 \int dn_3 \; P_A(a|n_2,n_3) \; P_B(b|n_1,n_3) \; P_C(c|n_1,n_2)$$

adapted from the locality constraint of the quantum triangle network

$$P_Q(a,b,c) = \int d\alpha \int d\beta \int d\gamma \; P_A(a|\beta,\gamma) \; P_B(b|\gamma,\alpha) \; P_C(c|\alpha,\beta)$$

We are constructing a Neural network that is able to approximate the distribution based on this form. Contrary to Bell's scenario we don't need to decide the measurements using an input parameter(the role of the Referee is not necessary here for nonlocality) A, B and C process their inputs with arbitrarily local response functions, and they each output a number out of four possible outcomes $a, b, c \in 0, 1, 2, 3$. Since what we have is one possible outcome out of four, the outputs of the perceptrons are conditional prob-

abilities $P_A(a|n_2, n_3), P_B(b|n_1, n_3)$ and $P_C(c|n_1, n_2)$ for a given input n_1, n_2 *and* n_3 i.e.three normalized vectors each of length 4.

We can analytically calculate the quantum distribution using the state we choose and the POVM measurement operators we choose to use. Since we have four choices for each of the three agents, we have the probability distribution $P(a, b, c)$ with 64 elements.

We can explicitly construct our quantum distribution $P_Q(a, b, c)$ by doing projective measurements on the state with the different measurement operators. Using this we can get an expression for the target probability distribution. We can take each source as the same entangled quantum state of two qubits. This helps in making the problem much more computationally tractable. Since we have the target distribution, we can now use the neural network to attempt to reproduce the target distribution, which results in the learned distribution.

After evaluating the Neural network for a batch of $n1, n2$, and $n3$ to approx the joint probability distribution $P(a, b, c)$ we use a Monte Carlo approximation, where we average over the product of probabilities over all the batches. We can get better results by increasing the number of batches.

$$P_M(a, b, c) = \frac{1}{N_{batch}} \sum_{i=1}^{N_{batch}} P_A(a|\beta_i, \gamma_i) P_B(b|\gamma_i, \alpha_i) P_C(c|\alpha_i, \beta_i) \tag{5}$$

The loss function can be any differentiable measure of discrepancy between the target distribution P_t and the neural network's output P_m(the learned distribution) We can use the Kullback-Leibler divergence

$$L(P_m) = \sum_{a,b,c} P_t(a, b, c) log(\frac{P_t(a, b, c)}{P_m(a, b, c)}) \tag{6}$$

Given a target distribution P_t the neural network trains and creates an explicit model for the distribution P_m which is closest to P_t. The learned distribution P_m is guaranteed to be from a local set by contradiction. If $P_t \approx P_m$ it's local and if not it's nonlocal.

There are 2 methods for analyzing the difference between the two distributions, we can vary the visibility from $v = 0$ to $v = 1$ i.e. decreasing the noise on a family of target distributions $P_t(v)$ by taking a distribution that is nonlocal and adding some noise controlled by the v parameter. So accordingly the $P_t(v = 0)$ will be a completely noisy(local) distribution and $P_t(v = 1)$ is the noiseless nonlocal distribution

By adding noise at some parameter of v^* we will enter the local set and stay in at $v < v^*$. We retrain the neural network for each noisy distribution and obtain a family of learned distributions $P_m(v)$. Observing

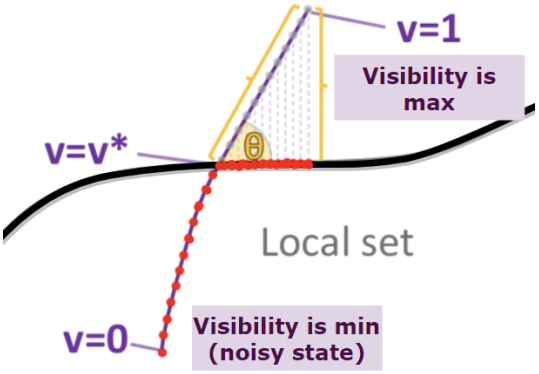


Figure 4: Geometry of local set when evaluating neural networks

a qualitative change at some point is an indication of traversing the local set's boundary. We can see this in the local response function of Alice, Bob, and Charlie.

$$d(P_t, P_m) = \sqrt{\sum_{a,b,c} [P_t(a,b,c) - P_m(a,b,c)]^2} \tag{7}$$

The other way is to find how far the learned distribution is from the target distribution

We can observe a clear lift-off at some point signaling that we are leaving the local set. We can also find the v^* and θ at which the learned distribution leaves the local set letting us have a quantitative study using some mild assumptions.

2.0.1 Fritz Distribution

We use the Fritz distribution[5] to benchmark the method we are using. Here the Bell scenario is wrapped in a triangle topology. Alice and Bob ρ_{AB} share a singlet, while Bob and Charlie ρ_{BC} and Alice and Charlie ρ_{AC} is either a maximally entangled or classically correlated state. Alice measures the ρ_{AC} in the computational basis, and based on that bit Alice then does a Pauli X or Z on ρ_{AB}. Bob does the same using Pauli $(X+Z)/\sqrt{2}$ and $(X-Z)/\sqrt{2}$ on ρ_{AB}. Both Alice and Bob give the measurement outcome and bit. Charlie measures both ρ_{AC} and ρ_{BC} in computational basis and gives 2 bits.

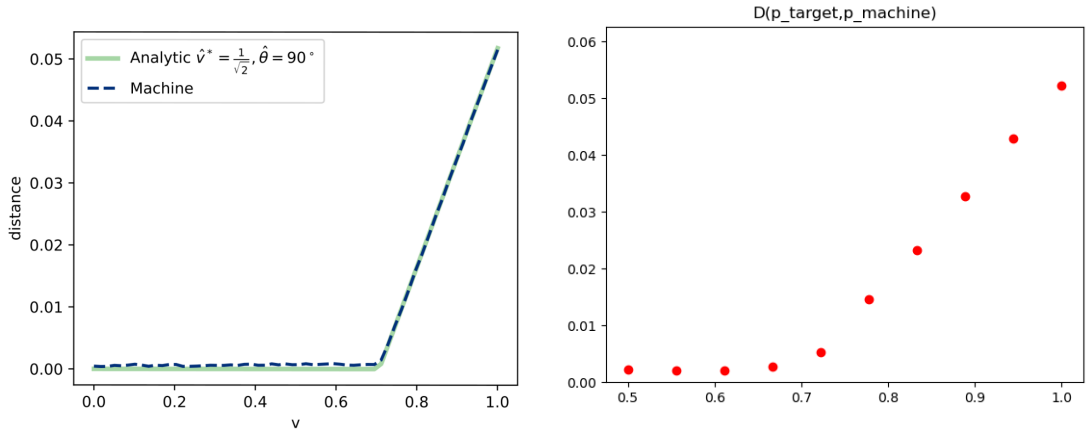
For the noise model, we can introduce finite visibility for the singlet shared between A and B, using the Werner state; where v is the visibility

$$\rho(v) = v|\psi^-\rangle\langle\psi^-| + (1-v)I/4 \tag{8}$$

for such a state we can see a clear lift off at $v^* = 1/\sqrt{2}$ and $\theta = 90°$

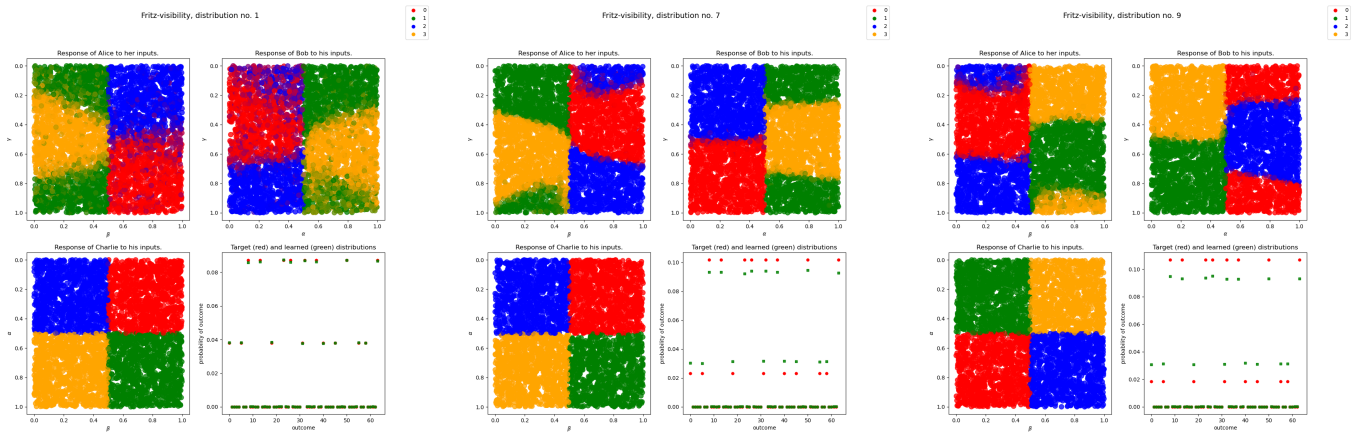
Response functions don't change much after v^*.This is because the machine finds same distribution for the nonlocal distributions that lie outside the local set.

This is due to the special case of Bell nonlocality and Fritz[2] where the local set is a polytope. Figure 5 shows how it will behave.(4)



(a) [2] (b)

Figure 5: Plot of euclidean distance perceived by the machine, $d_M(v)$ and the analytic distance $\hat{d}$(v) for $\hat{v}^* = 1/\sqrt{x}$ and $\theta = 90°$

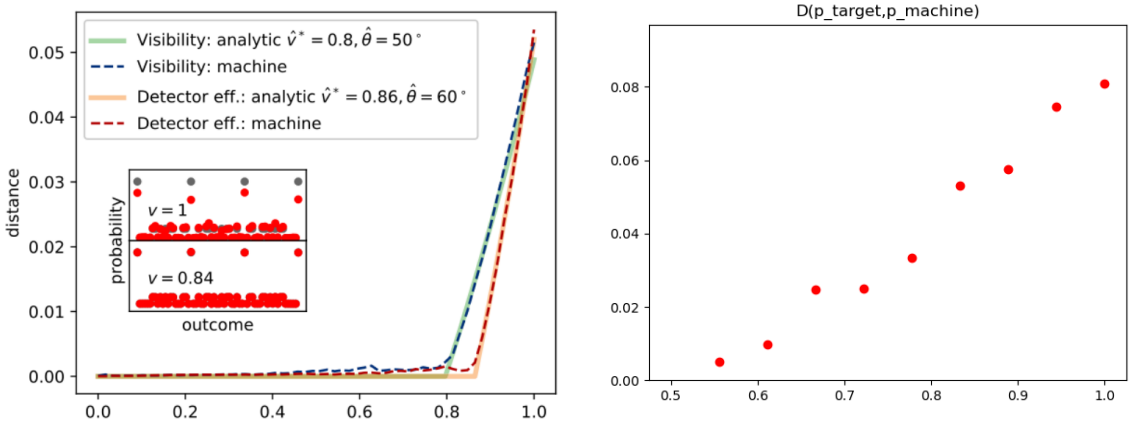


(a) v=0.1 (b) v=0.8 (c) v=1.0

Figure 6: Visualization of response functions and distance between p_t & p_m for different noise parameters

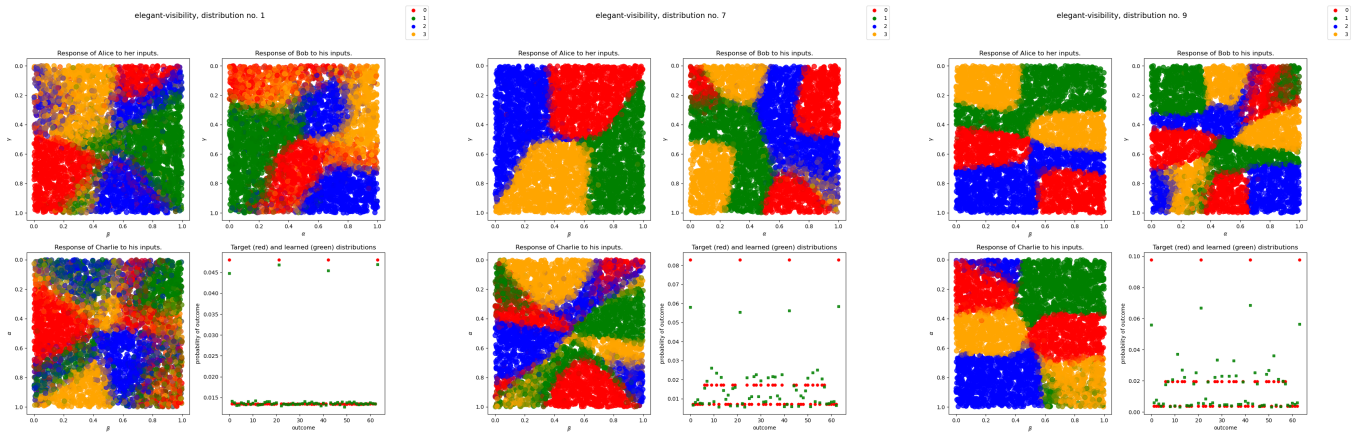
2.0.2 Elegant Distribution

This distribution[6] is much closer to the triangle structure as it combines both entangled states and entangled measurements. We introduce visibility to singlet sources such that all three have the form $\rho(v) = v|\psi^-\rangle\langle\psi^-| + (1-v)I/4$; where v is the visibility. Then for $v^* = 0.80$ and $\theta = 50°$ we can see a clear lift off.



(a) [2] (b)

Figure 7: Plot of euclidean distance perceived by the machine, $d_M(v)$ and the analytic distance $\hat{d}$(v)



(a) v=0.1 (b) v=0.8 (c) v=1.0

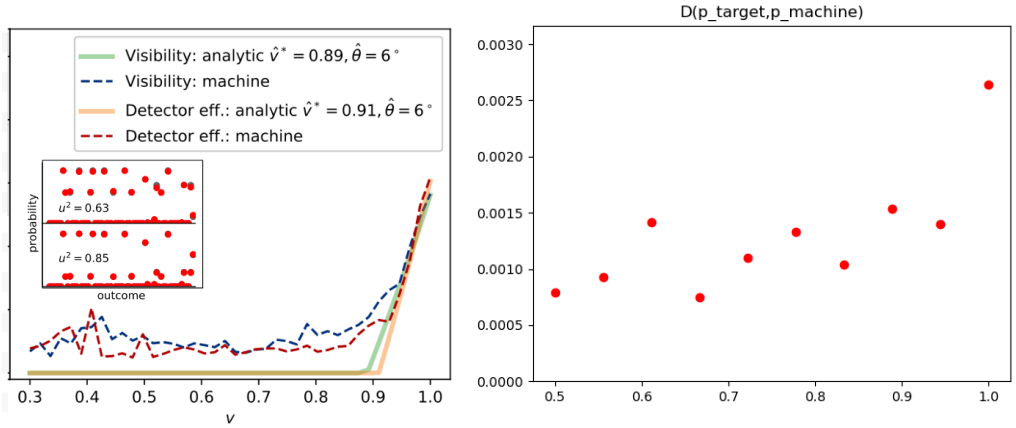
Figure 8: Visualization of response functions and distance between p_t & p_m for different noise parameters

2.0.3 Renou et al. distribution

To generate this distribution all three shared states are taken to be $|\psi^+\rangle$. Each party is performing the same measurement characterized by a single parameter $u \in [1/\sqrt{2}, 1]$ with eigenstates $|01\rangle, |10\rangle, u|00\rangle + \sqrt{1-u^2}|11\rangle, \sqrt{1-u^2}|00\rangle - u|11\rangle$

The authors[2] prove that the distribution is nonlocal $0.785 < u^2 < 1$ and that there exist local models for $u^2 \in 0.5, u^2_{max}, 1$

Looking at the noise robustness of distribution with $u^2 = 0.85$, which is approximately the most distant in the provenly nonlocal regime. We get to see clear lift off at visibility $v^* = 0.89$ & $\theta = 6°$
But the estimates are cruder than those of elegant due to target distributions being closer to the local set and the neural network must be getting stuck in local optima.



(a) [2] (b)

Figure 9: Plot of euclidean distance perceived by the machine $d_M(v)$ and the analytic distance $\hat{d}$(v)

3 Generalizing X states

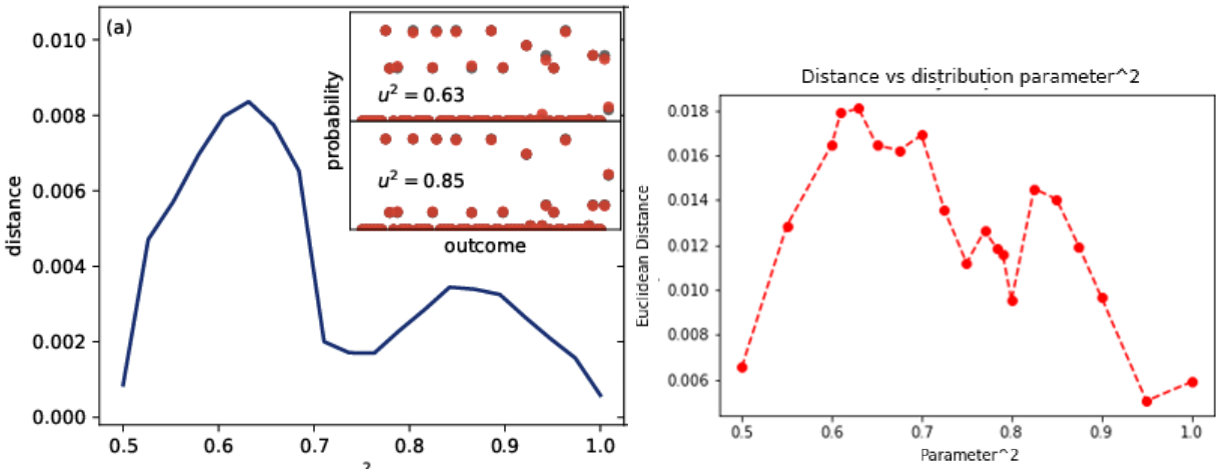
3.0.1 Classifying nonlocal distributions in Bell and Werner States using entanglement measurements

On using the Bell state $|\psi\rangle_{AB} := \frac{1}{\sqrt{2}}(|00\rangle_{AB} + |11\rangle_{AB})$ and an entangled set of measurements like [1]:

$$\begin{matrix} u|00\rangle + [\sqrt{1-u^2}]|11\rangle \\ [\sqrt{1-u^2}]|00\rangle - u|11\rangle \end{matrix} \quad \& \quad \begin{matrix} |01\rangle \\ |10\rangle \end{matrix}$$

These measurement operators only have the parameter u to have for over the entanglement measurements, After training the model and comparing the euclidean distance between the target and learned distributions, we got the following results.

- The work on triangle quantum networks by [1] showed that $P_Q(abc)$ is nonlocal for $0.785 < u^2 < 1$. They also proved that there exists local models for $u^2 \epsilon 0.5, 0.785, 1$.
- Later on examining the network structure using neural network modeling [2], nonlocal distributions were found outside the limit $0.785 < u^2 < 1$ where the best values where found at $u^2 = 0.63 \& 0.85$. The nonlocality peak at 0.63 was unexpected.
- The Neural network structure was able to successfully reproduce the nonlocal distributions $P_Q(abc)$ for $0.785 < u^2 < 1$ as well as the local distributions at $u^2 \epsilon 0.5, 0.785, 1$.



(a) [2] (b)

Figure 10: Plot of the distance perceived by the machine, $d_M(v)$ for different values of the entanglement parameter)

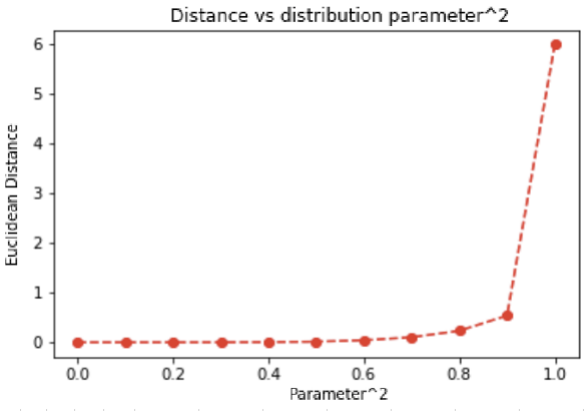


Figure 11: Noisy Werner state

- Using Werner states $\rho(v) = v|\psi^-\rangle\langle\psi^-| + (1-v)I/4$, we were able to get a control over the nonlocality of the distribitions by varying the noise parameter in the states.
- In the Standard Bell scenario, Werner states exhibit nonlocal behavior in the range of $1/\sqrt{2} < v <= 1$, where v is the visibility. Here in the triangle network scenario we can see clear nonlocal behavior from 0.9 to 1.0 visibility.
- Using the noisy Werner states we were able to get a quantitative measure of the nonlocal behavior, this particularly helps in creating nonmaximally entangled states in the triangle network structure.

Generalizing to X States

- Since we initially took all the bipartite qubit sources as the same. This significantly reduced the complexity of the problem. Morevover all the previous studies focused on Bell states and included only one set of enetangled measurement.
- We are generalizing the study to a bigger set called set of X states. Instead of the standard bipartite density matrix that has 15 parameters, X states have 7 paprameters.
- We have also included two sets of entangled measurements instead of just one in the measurement statistics.

The X states has a total of 7 parameters

$$\begin{bmatrix} a_3 + b_3 + c_{33} + 1 & 0 & 0 & c_{11} - ic_{12} - ic_{21} - c_{22} \\ 0 & a_3 - b_3 - c_{33} + 1 & c_{11} + ic_{12} - ic_{21} + c_2 & 0 \\ 0 & c_{11} - ic_{12} + ic_{21} + c_{22} & -a_3 + b_3 - c_{33} + 1 & 0 \\ c_{11} + ic_{12} + ic_{21} - c_{22} & 0 & 0 & -a_3 - b_3 + c_{33} + 1 \end{bmatrix}$$

Of the X states the Bell states are a small subset of it

The state we have studied is a type of X state with two parameters $\begin{bmatrix} a & 0 & 0 & a \\ 0 & b & b & 0 \\ 0 & b & b & 0 \\ a & 0 & 0 & a \end{bmatrix}$

And we have also included a two parameter entanglement measurement, for varying the degree of each of the two possible entangled measurements.

The Measurement basis are

$$\begin{matrix} u|00\rangle + [\sqrt{1-u^2}]|11\rangle \\ [\sqrt{1-u^2}]|00\rangle - u|11\rangle \end{matrix} \quad \& \quad \begin{matrix} w|01\rangle + \sqrt{1-w^2}|10\rangle \\ \sqrt{1-w^2}|01\rangle - w|10\rangle \end{matrix}$$

3.0.2 Optimization and Machine learning protocol

- We had used values from three independent uniform distributions from 0 to 1 as inputs to generate an array of 12 values for each batch. These 12 values contain the conditional probabilites of the the three agents A, B and C. On multiplying them selectively we can get the 64 element probability distribution of the quantum distribution in the 64 dimensional Hilbert space.
- Like in [2] we had used the Kullback-Leibler divergence for the Loss function and Monte Carlo approximation for calculating the distribution.
- Since Non-Bell states turned out to be difficult to train, we had to include multiple sweeps with different variations of the optimizing algorithm and learning rate to find better optimas.
- We also had success using evolutionary algorithm such as Genetic algorithms to optimize the training, but since we are dealing with a huge 64 dimensional Hilbert space the algorithm was slow at converging to an optimal distribution compared to the standard adadelta optimization.
- We made use of the Padmanabha computational cluster made available through the center for High Performance Computation at IISER-TVM for the parallel training of the neural network models

To increase the accuracy of the results we had increased the batch size of the training inputs and the number of parameters in the Neural network architecture. We have three perceptron neural networks connected by the locality constraint summing up to give a total of 22,230 parameters.

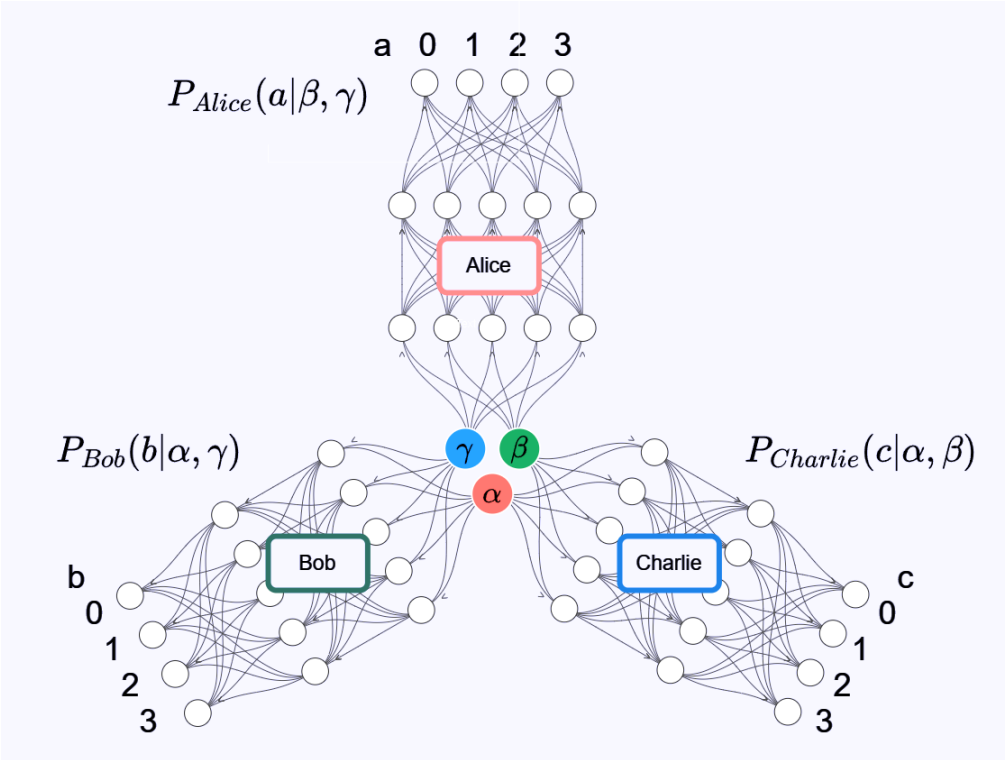


Figure 12: Neural network architecture

4 Results and discussion

3.0.1 Best measurement settings for X States

$$|\psi\rangle_{AB} := \frac{1}{\sqrt{2}}(|00\rangle_{AB} + |11\rangle_{AB}) \quad |\psi\rangle_{AB} := \frac{1}{\sqrt{2}}(|01\rangle_{AB} + |10\rangle_{AB})$$
$$\&$$
$$|\psi\rangle_{AB} := \frac{1}{\sqrt{2}}(|00\rangle_{AB} - |11\rangle_{AB}) \quad |\psi\rangle_{AB} := \frac{1}{\sqrt{2}}(|01\rangle_{AB} - |10\rangle_{AB})$$

using only one set of entangled measurements

$$u|00\rangle + [\sqrt{1-u^2}]|11\rangle \quad w|01\rangle + \sqrt{1-w^2}|10\rangle$$
$$\&$$
$$[\sqrt{1-u^2}]|00\rangle - u|11\rangle \quad \sqrt{1-w^2}|01\rangle - w|10\rangle$$

The measurement settings were found to be 0.63 and 0.85 and we were able to reproduce the results by Renou [1] and Krivachy [2].

We can use our input state as a combination of two bell states
$|\psi\rangle_{AB} := \frac{1}{\sqrt{2}}(|00\rangle_{AB} + |11\rangle_{AB})$ and $|\psi\rangle_{AB} := \frac{1}{\sqrt{2}}(|01\rangle_{AB} + |10\rangle_{AB})$

to create a X state with 2 parameters, which should also satisfy the properties of density matrices to represnt a quantum state.

$$\begin{pmatrix} a & 0 & 0 & a \\ 0 & b & b & 0 \\ 0 & b & b & 0 \\ a & 0 & 0 & a \end{pmatrix}$$

We can then assign X states using these two parameters to study their nonlocal distributions. We are also using two sets of entangled measurements

$$\begin{matrix} u|00\rangle + [\sqrt{1-u^2}]|11\rangle \\ [\sqrt{1-u^2}]|00\rangle - u|11\rangle \end{matrix} \quad \& \quad \begin{matrix} w|01\rangle + \sqrt{1-w^2}|10\rangle \\ \sqrt{1-w^2}|01\rangle - w|10\rangle \end{matrix}$$

- For $a = 1$ & $b = 0.5$ and $a = 0.5$ & $b = 1$ we retain the euclidean distance peaks vs u parameter at $u^2 = 0.63$ & 0.85 and we can see the emergence two smaller peaks in between these. We could also retain the local points at $u^2 \epsilon \{0.5, 0.785, 1\}$.

- We are generalizing the study to a bigger set called set of X states. Instead of the standard bipartite density matrix with 15 parameters, X states have 7.

- On expanding to a more general set of two parameter X states (a & b that follow the density operator conditions) with 2 parameter joint ent meas ($u^2 \& w^2$) We have found two distinct peaks at (0.4 0.8) & (0.8 0.5). Since the four measurement operators have changed, the peaks have readjusted from 0.63 and 0.85 to these points.

- The two best choices are symmetric to the u = w line. This is the most interesting part, in that these measurements are symmetric and the two best measurements are related by swapping each other.

- For X states the best measurement choices have shifted towards $u^2 = 0.5 \& w^2 = 0.5$. The best measurement choices shift as we go changing the parameters of the X states

- This suggests that there is some global property. As for this behavior we observed that when we use bell states we get the best peaks at $(u^2, w^2) = (0.5, 0.95)$ & (0.95,0.5), but as we include another bell states in linear combination the peaks gradually shift to (0.5,0.5) averaging around (0.8,0.5) & (0.5,0.8). This might be due to the new correlations that occur due to the new basis vectors introduced when combining Bell states

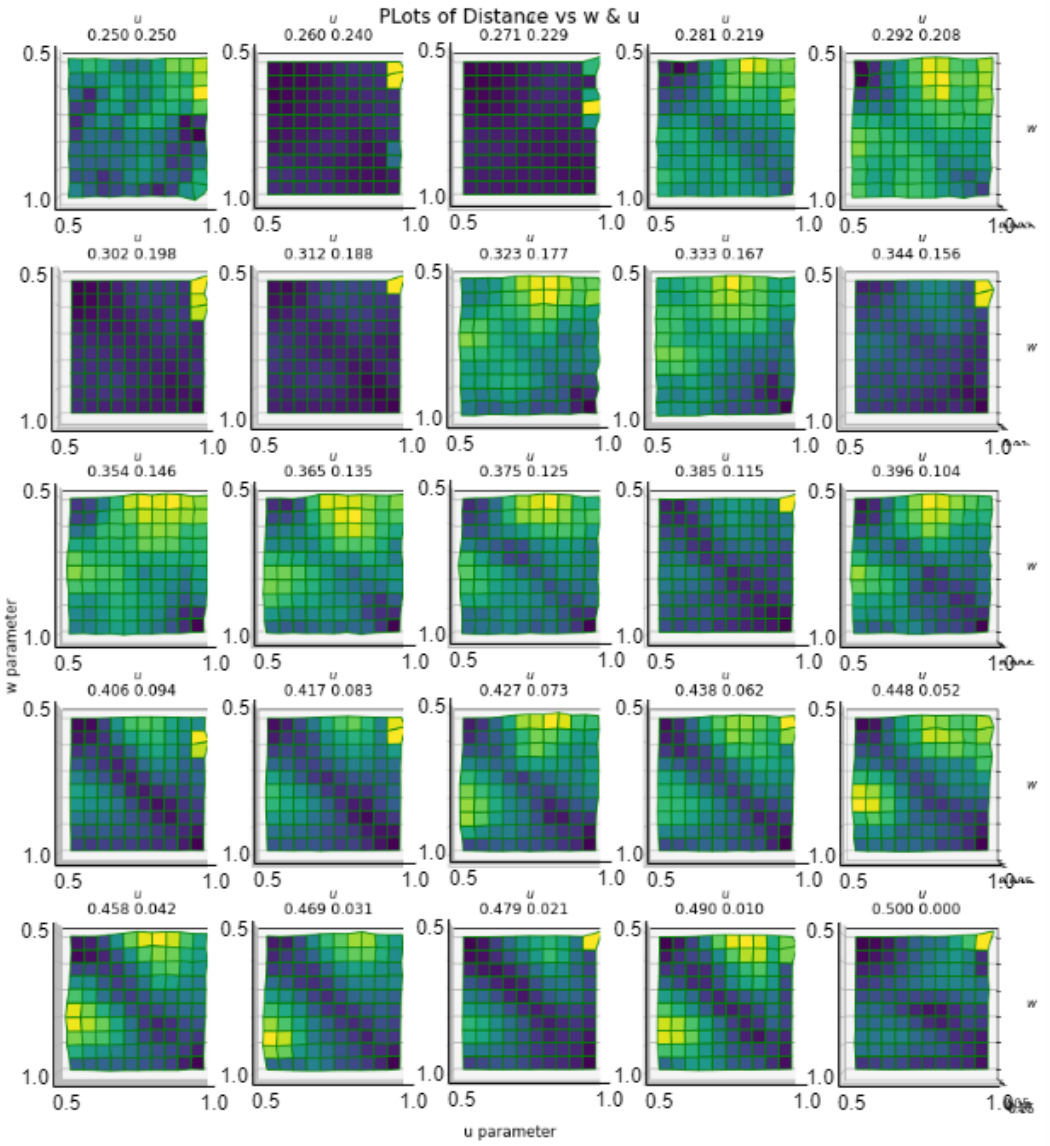


Figure 13: LHV Model euclidean distance peaks

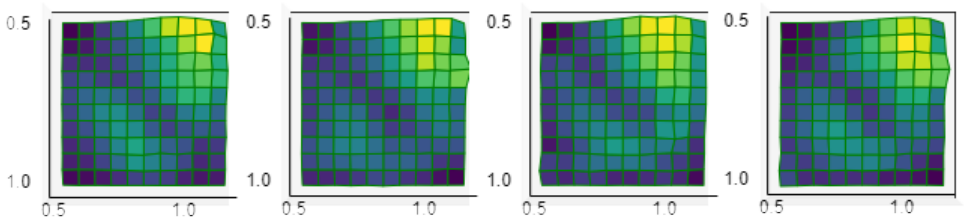
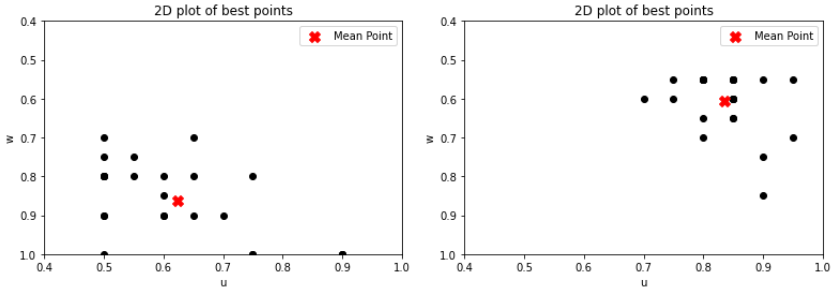


Figure 14: Bell state best measurement settings



(a) (b)

Figure 15: The first figure focues on the points of the peak towards the left bottom (0.8,0.5), and the other graph to peak at the right top (0.5,0.8)

- We have identified that the region around $(u^2, w^2) = (1.0,1.0)$, is almost devoid of any peaks and is fully local as described by Renou. We also find that the distribution with u = w are also local.
- Furthermore we have found from the analytical equations that when we change u & w we will get the same distribution if we change a & b as well this was reflected in our data. Also it was confirmed the a & b values being the same will be giving local distributions and the difference between the two being the maximum is ideal for the best nonlocal distribution we can observe

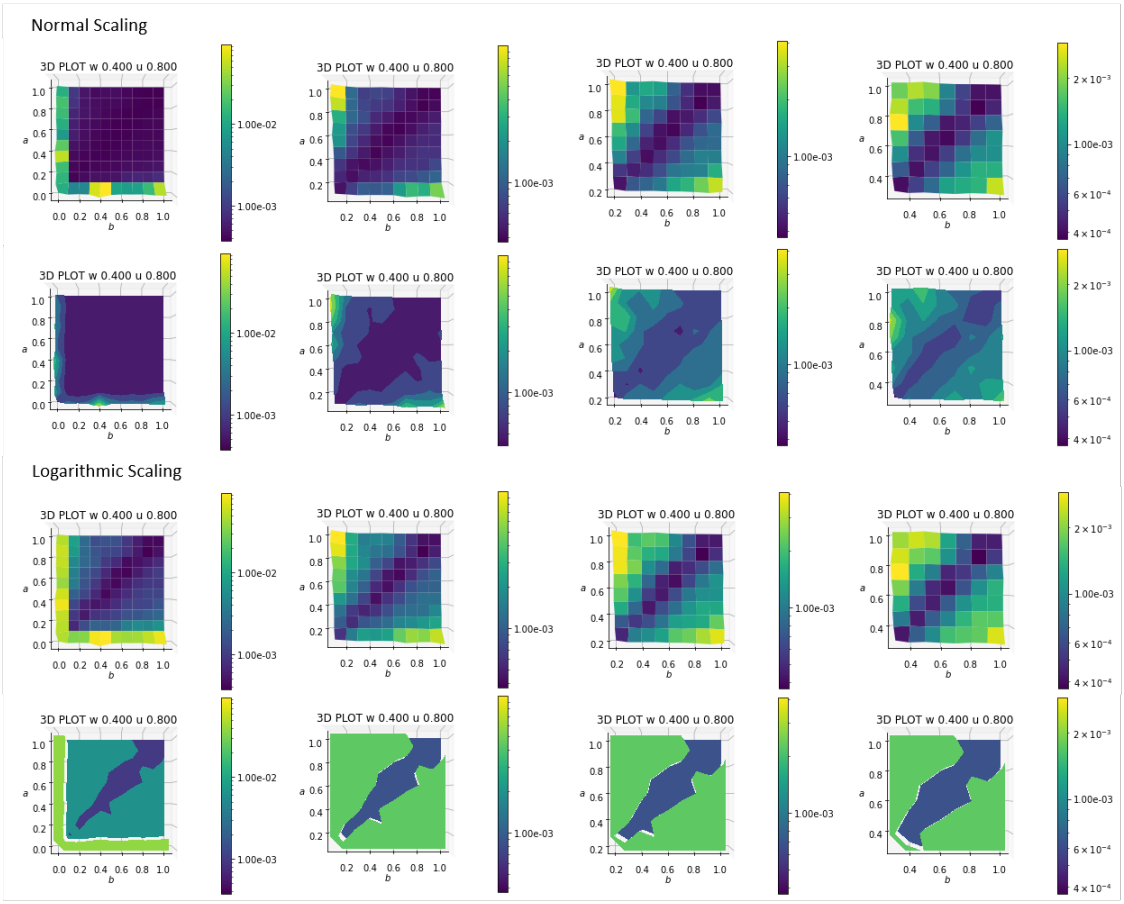


Figure 16: The euclidean distance graphs for change in quantum state parameters with the suitable fixed measurement(0.50,0.80) for both Normal and Logarithmic Scaling

We have a peculiar valley in between the two peaks and the symmetry in the graph is because the probability distribution elements get symmetry from the symmetry of our density matrix (a & b). We believe that the non-Bell state's best peaks are aligning closer to each other due to the correlation between states of (00,11) & (01,10) basis

These are the euclidean distance plots for (u,w) = (0.5,0.8) & (0.8,0.5). The left one has the full parameter space of (a,b) from 0 to 1, in the right one the a = b = 0 parts were omitted. As you can see the left parts have the Euclidean distances (a measure for nonlocality in the network) really high almost

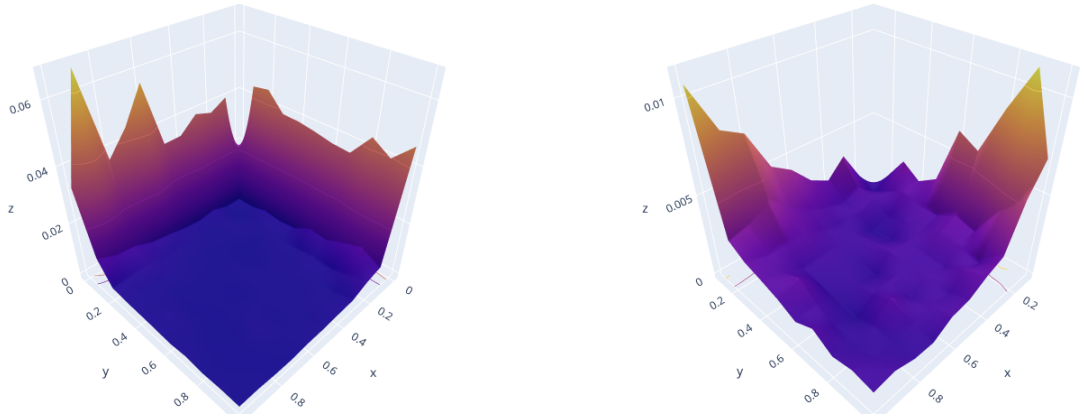


Figure 17: Measurement case for full set of quantum state parameters, and excluding extreme points

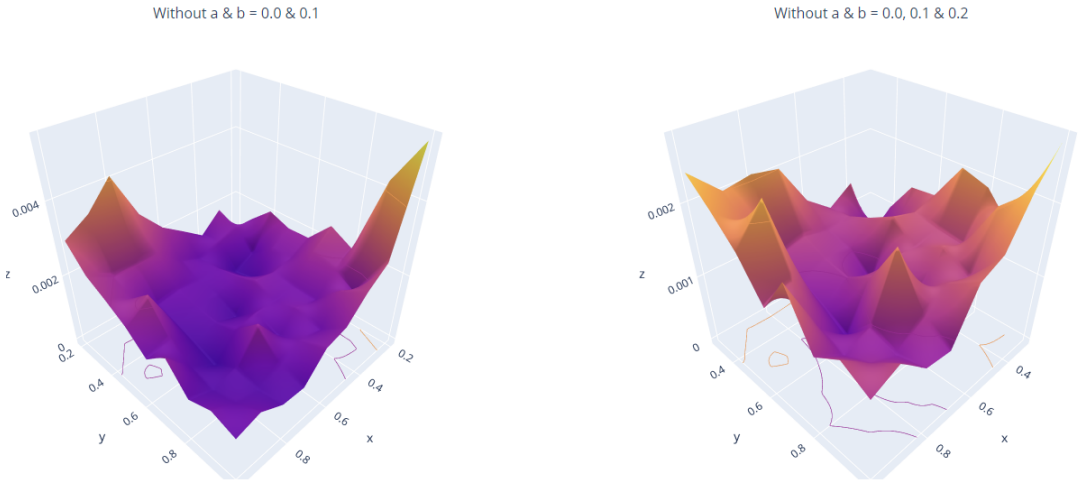
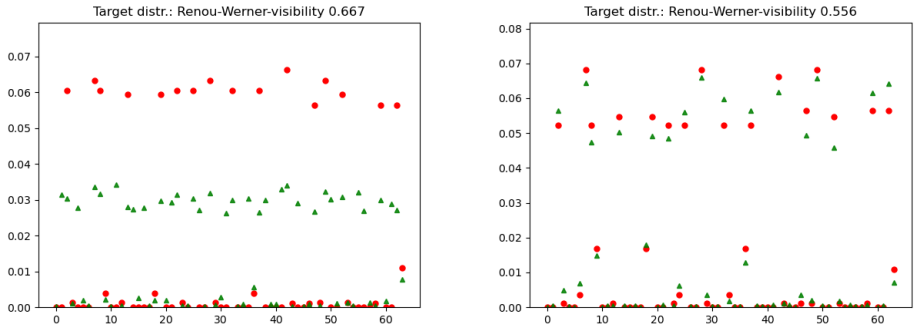


Figure 18: Measurement case for quantum states

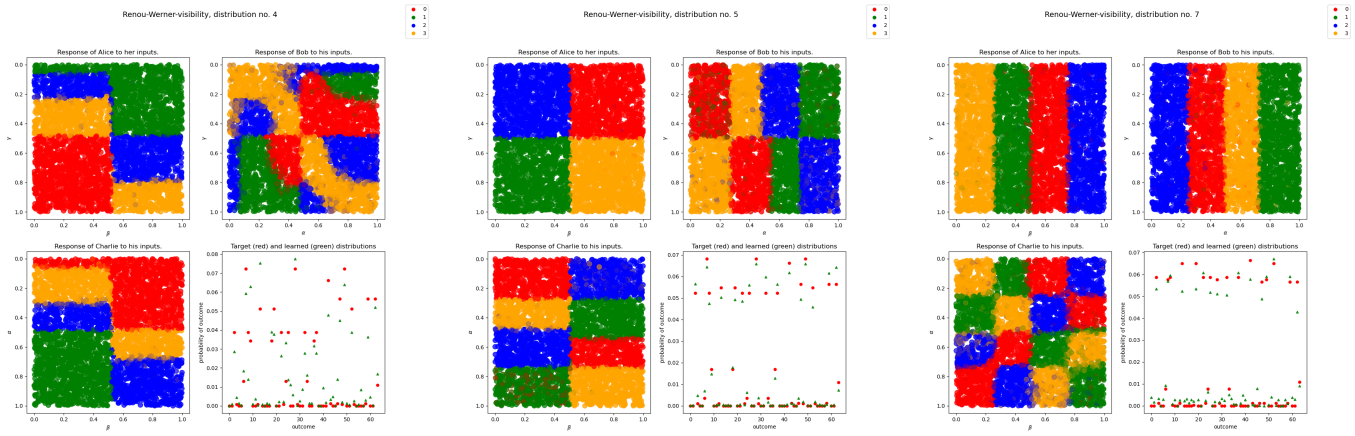
like a wall, this is due to it being from the Bell states. The right part has peaks that are smaller because we have removed the Bell state parts alone, this lets us focus on the X states, here we can see that a = b are local points and they form the trough in the middle, towards the sides we can see the peaks gradually increasing to where a is most greater than b and viceversa

The behavior of symmetry between the two measurement settings could result from some global property of the nonlocal and local distribution boundary in the triangle network scenario. Understanding such global properties will give us a direction towards understanding the nonlocal and local space of this complex problem.



(a) (b)

Figure 19: The distributions show that only a few elements of the 64 probability distributions are responsible for nonlocal behavior



(a) (b) (c)

Figure 20: We also got interesting symmetric behavior in the response functions of these distributions

From the Euclidean distances, we can also identify that only certain elements of the 64-element distribution are different; only certain subspaces are responsible for the nonlocal behavior exhibited by the triangle quantum network. Also since we can use the same set of measurement operators for the agents we end getting the same distribution for many of the possibilities.

4 Conclusion and future direction

- We were able to reproduce the results of Renou [1] and Krivachy [2], by using the neural network architecture to classify nonlocal states. We also got the best parameter for entangled measurements when considering Bell states and then expanded it to X states.
- We discovered a symmetry in the measurement setting, the significance of which is a work in progress. And new results will be soon available.

- We were able to control the nonlocality measure, by making use of the noise parameter in Werner state; we also got an estimate of the visibility required for a Bell state to show nonlocal behavior.
- For Renou[1] distribution the distance perceived by the machine $d_M(v)$ was not optimized, because the nonlocal distribution lies very close to the local space. We improved the optimization by increasing the neural network parameters and using multiple sweeps with varying optimizers to get an optimal result. And we were able to reproduce the genuine triangle nonlocality results [1].
- We were able to expand upon the study on Bell states to X states; we also included two sets of entangled measurements. For a two-parameter set of X states we were able to identify the best possible entanglement measurement operators for generating a nonlocal distribution.
- We also identified global properties in the behavior of nonlocality for different measurement operators with different states. We speculate that these properties will explain how we can better understand and navigate the boundary between local and nonlocal distribution space.

3.0.1 Future direction

- Expand on the X states and study on the full set of 7 parameters to find the best measurement setting for all X states for showing nonlocal behavior.
- Understand the behavior of the nonlocality euclidean distance measure for different states and different entanglement measures, and work towards understanding and navigating the nonlocal and local space for the traingle quantum network.
- Since the neural network acts as a classifier, modify the condition to select distributions that are x distance away from the local space. Since we have encountered interesting behaviors [1] close to the local space, studying other similiar distributions can give us important insights.
- Improve the optimization and adjust the machine learning parameters to prevent the neural network from getting stuck in local optima to better distinguish between the learned and target distribution.
- Find a measure for nonlocality in the triangle setting independent of a trained model.
- Use other machine learning approaches such as reinforcement learning to train the neural network to navigate the local and nonlocal distribution space.
- Continue the study on noise robustness[2], using the distance and the response functions variations; this can further help to map the nonlocal boundary.

- Build neural network structures for different distributions and find a structure that works for a broad range of systems

By understanding such correlations that cannot be described by a casual model, we can use them in quantum information processing, reducing communication complexity, QKD, private randomness generators, device-independent entanglement witnesses, etc. The triangle network is especially useful as it gives nonlocality independent of the measurement we do. These correlations also help in understanding the fundamental aspects of quantum theory and quantum information.
Finding non-maximally entangled nonlocal states are particularly useful because in experimental conditions its very difficult to create and maintain maximally entangled states.

We thank the use of the Padmanabha computational cluster which was made available through the center for High Performance Computation at IISER-Thiruvananthapuram.

References

[1] Marc-Olivier Renou et al. "Genuine Quantum Nonlocality in the Triangle Network". In: *Physical Review Letters* 123.14 (Sept. 2019). DOI: 10.1103/physrevlett.123.140401.

[2] Tamás Kriváchy et al. "A neural network oracle for quantum nonlocality problems in networks". In: *npj Quantum Information* 6.1 (Aug. 2020). DOI: 10.1038/s41534-020-00305-x.

[3] J. S. Bell. "On the Einstein Podolsky Rosen paradox". In: *Physics Physique Fizika* 1 (3 Nov. 1964), pp. 195–200. DOI: 10.1103/PhysicsPhysiqueFizika.1.195.

[4] Nicolas Brunner et al. "Bell nonlocality". In: *Reviews of Modern Physics* 86.2 (Apr. 2014), pp. 419–478. DOI: 10.1103/revmodphys.86.419.

[5] Tobias Fritz. "Beyond Bell's theorem: correlation scenarios". In: *New Journal of Physics* 14.10 (Oct. 2012), p. 103001. DOI: 10.1088/1367-2630/14/10/103001.

[6] Nicolas Gisin. "Entanglement 25 Years after Quantum Teleportation: Testing Joint Measurements in Quantum Networks". In: *Entropy* 21.3 (Mar. 2019), p. 325. DOI: 10.3390/e21030325.

[7] Cyril Branciard et al. "Bilocal versus nonbilocal correlations in entanglement-swapping experiments". In: *Physical Review A* 85.3 (Mar. 2012). DOI: 10.1103/physreva.85.032119.

[8] Ryszard Horodecki et al. "Quantum entanglement". In: *Rev. Mod. Phys.* 81 (2 June 2009), pp. 865–942. DOI: 10.1103/RevModPhys.81.865.