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

December 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

Acknowledgments

This dissertation has been a great learning experience for me and I thank Prof. Anil Shaji for his astute guidance and deep insights that have helped build my understanding of the subject and importantly my original research. Dr. Debashis Saha has been a constant support and mentor in this process, and I am thankful to him for proposing the interesting project on "Genuine Quantum Network nonlocality, for his guidance and inputs for our research paper and my Masters' thesis. I am thankful to them for the long and absorbing discussions about quantum foundations, directions, and critical insights for our work. The members of the Quantum Information Theory group have been a constant support throughout the group meetings, discussions in keeping me up to date with the latest research and with my personal queries.

All this was also supported by the Chanakya Fellowship, I-HUB QTF National Mission on Interdisciplinary Cyber Physical Systems, Department of Science and Technology, Government of India. I also thank the use of the Padmanabha computational cluster which was made available through the center for High Performance Computation at IISER-Thiruvananthapuram, Kerala.

Thanks are due to Arunima for all her support and for accepting to be the first audience for many of my discussions and musings. Lastly, but importantly, my family, has been the inspiration behind my academic pursuits, impressing upon me at a young age that knowledge should be sought not merely as means to an end, but as an end in itself.

Thiruvananthapuram - 695 551 **Anantha Krishnan S**

April 2023

ABSTRACT

Two projects are part of this Masters thesis

Quantum Networks, GNNL[1] and LHV[2] Neural Networks: This is my own journey through this domain, understanding and exploring the various intricacies that quantum networks give rise to that are disjoint from an origin in Standard Bell Nonlocality. This is also where I realised that designing machines and methods combining Analytical methodologies and Machine Learning numerics holistically help us view and study these systems with greater depths of understanding. An abstract of this provided in the preface of this report.

The Best Measurement Settings and Quantum Resources for Witnessing Genuine Nonlocality in the Triangle Quantum Network: The research carried out as part of the thesis on Quantum Networks lead to my own contribution in this domain. This was an exciting experience to learn and apply my experience in Quantum Information theory and Machine Learning. Here is an abstract of the part of this masters thesis which is currently in works of being published (check here).

Nonlocal correlations lie at the core of quantum foundations, and distributions displaying them are of great importance in applications of information processing. Unlike quantum systems with a convex local boundary, the nonlocality without inputs in networks is of particular interest as it deviates from the standard bell inequality violation criteria. Nonlocality being a correlation that can't be reproduced by a classical structure, machine learning with classical sources was later introduced as a counter test for the same. These works were particular to bell states; we are expanding on this work with X states and joint entangled measurements. Consequently, we have observed the measurement choices for X states with two-parameter entangled Bell state measurements, which can be almost confidently extrapolated to all two qubit quantum sources capable of showing recognizable nonlocal correlations. Then we further categorized the states showing higher nonlocal correlations and their respective measurement range. We backed up our results with noise robustness of these nonlocal distributions and their further analyses. We also came across and further explored whether a triangle network can produce nonlocal correlation with just entangled measurements on classically correlated states and analytically proved that we cannot. We then used our analytical insights to model a new LHV neural network model capable of not just distinguishing pure states but all non-separble nonlocal distributions with quantum state density matrices of rank k, with increasing model complexity of the order k. Exploring this new domain of nonlocality will give us glimpses into the defining principles of quantum nonlocality and its wide applications.

[1] Genuine Network Nonlocality
[2] Local Hidden Variable

Preface

Here we provide a preface of the review part of this masters thesis on Quantum Networks. This is commendatory to the abstract on my original contribution on "The Best Measurements Settings for Witnessing Genuine Nonlocality in the Triangle Quantum Network". This primarily begins with a review of the different Quantum Triangle Network distributions including the Bilocality Network[1] and the RGB4[1][2] distribution which I focused on for my research work. After exploring the joint entangled measurements and the best measurement settings we found for expressing GNNL[2] in the Quantum Triangle Network, we move on to an exploration of the quantum resources required for the same and how it brings in different characteristics into the picture. We later explore if entangled local measurements are enough to express GNNL[2] and prove analytically that this isn't the case.

We have complemented our research with both numerics and analytics as I believe doing so lets you see these systems and theories with a greater depth of understanding. Moreover I really enjoy building machines capable of understanding and answering our foundational questions.

The study of quantum correlations is pivotal in understanding quantum nature both in foundational aspects of quantum theory and in advancing quantum information technologies. Originating from Bell's seminal work [3], with subsequent research leading to the theory of Bell nonlocality has been a cornerstone in this field. In the context of Bell scenarios, parties share classical or quantum sources, which upon measurement give rise to correlations classified as local or nonlocal, respectively. Investigating multipartite domains has further elucidated these correlations, particularly in network configurations where conventional Bell scenarios can be represented as elementary manifestations. Unlike standard setups, networks entail distant observers doing joint separable and entangled measurements sharing entanglement sourced from multiple, independent entities. Through these joint measurements on entangled states, observers establish strong correlations across the network, opening up novel questions and revealing new properties. Notably, such networks fundamentally can't be explained by locally causal models, inspired from this local hidden variable machine learning models can be used to learn and to discern quantum from classical correlations. These interdisciplinary approaches using both analytical and machine learning techniques has led to significant insights, including the discovery of novel quantum correlations within network structures[4]. Furthermore, employing classical machine learning to comprehend quantum phenomena sheds light on foundational aspects of quantum theory and information science. Understanding how classical machines learn about quantum structures is of paramount interest, offering avenues for uncovering new quantum phenomena and designing improved methodologies for their exploration. This underscores the pivotal role of machine learning, both classical and quantum, in deciphering the intricacies of quantum systems, akin to its efficacy in addressing classical problems.

[1] Genuine quantum nonlocality in the triangle network
[2] Genuine Network Nonlocality

Contents

List of Figures

Introduction

Structure of this thesis

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 decentralized network structures, in which several independent sources are shared among the parties over a network. We then describe the quantum distributions that networks can give rise to, including the renou et. al[2]. distribution that show nonlocality without inputs. These novel distributions arise due to the special entangled measurement play in Quantum networks; like the correlated sources the measurements are also important resources in networks. Hence, it can be seen throughout the report that these measurements are emphasized. 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. Here, we also point out how and why we understand these quantum systems using machines, why this work and many other works seed the aspiration in building quantum machines capable of understanding the mysteries of our universe.

We then elaborate on the methods part detailing how we can use LHV Neural Network Models in learning and classifying local and nonlocal distributions. We then generalize the study to X states and analyse the results. Based on the generalized cases we sampled across it is safe to extrapolate our results on the measurement settings to any two-qubit state capable of showing high nonlocal correlations. We also classified the quantum states capable of doing the same and confirmed it by further analyses. Apart from the results section we also came across an interesting ordeal or perhaps an interesting question on whether a triangle network can produce nonlocal correlation with just entangled measurements on classically correlated states, we further explored this and our current work encompasses concluding this, which is further discussed in the future directions and concluding remarks.

Quantum States, Measurements and Entanglement

Quantum states are identified with their finite or infinite Hilbert space H. Why H, because a d-dimensional pure quantum states can be mathematically described a vector living in the Hilbert space of $H = \mathbb{C}^d$ dimensions. For d = 2, we refer to these quantum states as qubit states and if the dimension is higher, as qudit states.

Then comes the mandatatory definitions of pure and mixed states for anyone who want to understand the subject. The pure state is a normalized vector and is called a ket vector $|\psi\rangle \in \mathbb{C}^d$. Its dual vector is called the bra vector $\langle\psi|$, where $\langle\psi| = |\psi\rangle^\dagger$. A pure state $|\psi\rangle$ can be expressed as a linear combination of some orthonormal basis $|a_i\rangle_{i=0}^{d-1}$:

$$|\psi\rangle = \sum_{i=0}^{d-1} c_i |a_i\rangle$$

where the complex-valued coefficients c_i satisfy the constraint $\sum_{i=0}^{d-1} |c_i|^2 = 1$, because of the normalization constraint, $\langle\psi|\psi\rangle = 1$.

If the basis $|a_i\rangle$ correspond to the standard basis in linear algebra, $|i\rangle_{i=0}^{d-1}$, they are called computational basis and each element of the basis set $|i\rangle$ is then a computational basis element. Whereas, the square of the absolute value of each coefficient c_i, corresponds to the probability of obtaining an outcome a_i, if the system was measured in the basis $|a_i\rangle_{i=0}^{d-1}$.

Often it happens that one does not have an exact information about a state of a quantum system, but rather knows that with some probability p_i the system is in a pure state $|\psi_i\rangle$. In this case entire physical system is in a mixed state and can be described by a density matrix ρ:

$$\rho = \sum_i p_i \psi_i \rangle\langle \psi_i|$$

Since p_i's are probabilities, that is $p_i \geq 0$ and $\sum_i p_i = 1$, it follows that $tr(\rho) = 1$. At the same time, each matrix $|\psi_i\rangle\langle\psi_i|$ is hermitian and positive. Hence, a density matrix ρ is a complex-valued trace-one matrix with non-negative real eigenvalues, $\rho \geq 0$. Moreover, any trace-one semi definite matrix is a density matrix describing a mixed state. From here it follows that the set of all quantum states is a closed convex set with pure states on its boundary.

Until now we regarded a quantum system as a single system living in a d-dimensional Hilbert space H^d. However, we can consider a systems containing N quantum states, each with a respective Hilbert space dimension, $d_1, ...d_N$. Then a pure state $|\psi\rangle$ is a state vector ina composite Hilbert space dimension, denoted by the tensor product $H^{d_1} \otimes H^{d_2} \otimes ...H^{d_N}$:

$$|\psi\rangle = \sum_{i_1,...i_N=0}^{(d_1-1),...,(d_N-1)} c_{i_1...i_N}|i_1\rangle \otimes ... \otimes |i_N\rangle = \sum_{i_1,...i_N=0}^{(d_1-1),...,(d_N-1)} c_{i_1...i_N}|i_1...i_N\rangle$$

, where the right hand side is a short hand notation, first, omitting the tensor product and then merging two ket vectors into one: $|a\rangle \otimes |b\rangle \equiv |a\rangle|b\rangle \equiv |ab\rangle$

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 being separated, a referee gives Alice and Bob a random bit x, y each of which decides the measurement outcome of Alice and Bob's different measurement options. After the measurement, the outcomes are a and b. 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 two systems that were established when they interacted in the past.

So, if we include all the past factors, the residual indeterminities about the output should be decoupled[5].

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

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 results $a, b \in -1, 1$, 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 $|\langle S\rangle| \leq 2$.

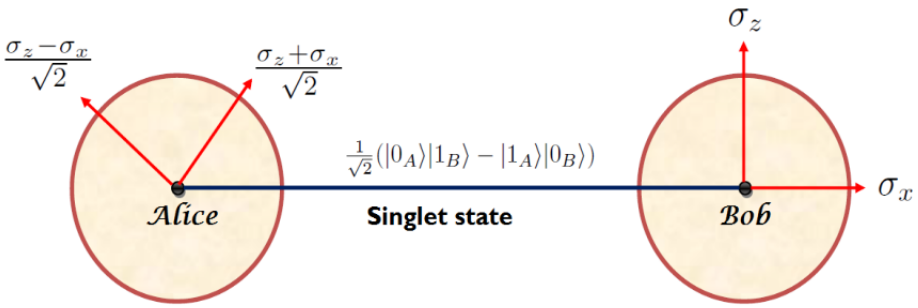


Figure 1: Alice's and Bob's measurement choice

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

$$|\langle S\rangle| = |\langle a_0 b_0\rangle + \langle a_0 b_1\rangle + \langle a_1 b_0\rangle + \langle a_1 b_1\rangle|$$

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

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.

Nonlocality in Quantum Networks

Recently, growing interest has been devoted to the exploration of quantum nonlocality in networks. A general framework has been developed for this problem. Consider a network with several parties, and N sources distributing physical (classical or quantum) systems to various subsets of parties. The main idea consists in assuming that all the sources of the network are independent from each other. This allows for the notion of N locality, which can be seen as a natural extension of the concept of Bell locality to networks with independent sources. Correlations that do not admit an N-local model are termed "network nonlocal." Importantly, whether a given distribution is network nonlocal or not depends on its quantum resources including entangled measurements and the topology of the considered network. In particular, for certain networks there are correlations that are provably network nonlocal, while the same correlation would be local when considering the usual Bell scenario.

Bilocality Network

Quantum theory allows for correlations in networks that are network nonlocal. The first examples of "quantum network nonlocality" have been derived for the scenario of entanglement swapping, referred to as "bilocality".

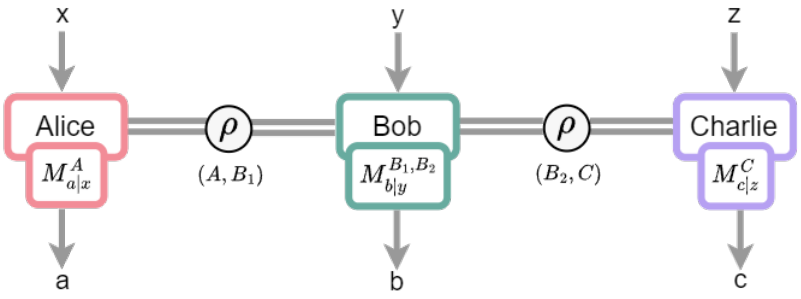


Figure 2: Bilocality Network: Entanglement Swapping

In turn, a more astonishing effect was discovered, namely, that quantum network nonlocality can be demonstrated without the need for measurement inputs, i.e., each party performs a single fixed measurment. This would of course be impossible in the standard Bell scenario, and hence appears at first sight as a novel form of quantum nonlocality proper to networks. A first example of such "quantum nonlocality without inputs" was proposed by Fritz, for the so called triangle network, it features three parties, each pair of them connected by a different source. After inspection, it turns out, however, that Fritz's example can be viewd as a clever mapping of the standard Bell Clauser-Horne-Shimony-Holt (CHSH) scenario into the triangle network configurration.

Triangle Network

A second example was more recently proposed by Renou et al, also for the triangle network. The authors argued, however, that their example is fundamentally different from that of Fritz and could represent a form a form of quantum nonlocality genuine to networks, i.e., that cannot be mapped back to any standard Bell scenario. This conjecture relies on the observation

that their quantum distribution was based on all three sources producing an entangled state, and all parties performing an entangled measurement on their incoming subsystems. That is 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.

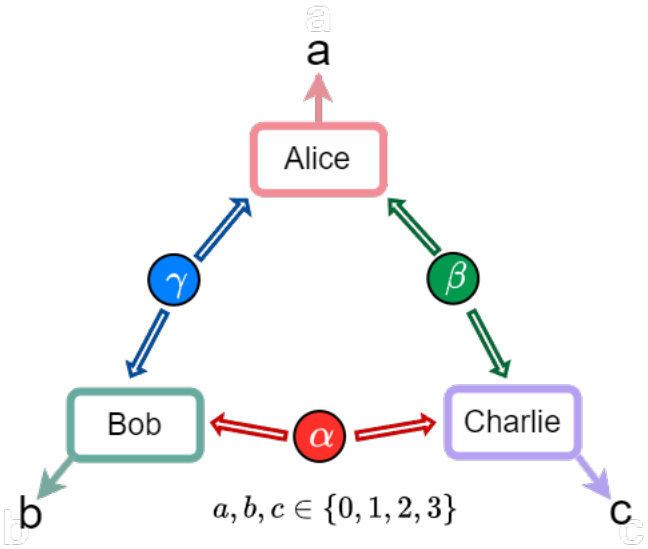


Figure 3: Triangle Causal Network

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 on which measurement setting to use. The statistics of the experiment are thus given by the joint probability distribution P(a, b, c).

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 act 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.

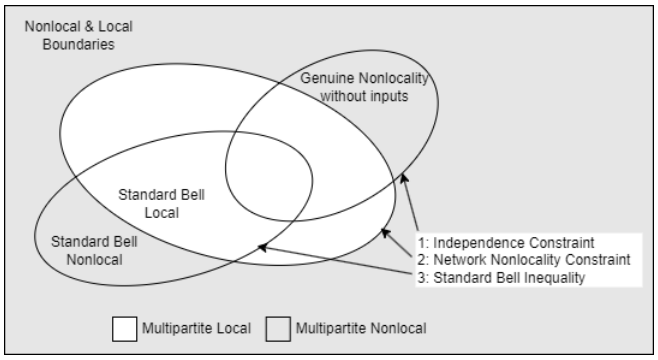


Figure 4: Local and Nonlocal domain

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 [2] 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 [6].

Machine learning

The rise of machine learning in recent times has remarkably transformed science and society. The goal of machine learning is to get computers to act without being explicitly programmed. Some of the typical applications of ma- chine learning are self-driving cars, efficient web search, improved speech recognition, enhanced understanding of the human genome and online fraud detection. This viral spread in interest has exploded to various areas of science and engineering, in part due to the hope that artificial intelligence may supplement human intelligence to understand some of the deep problems in science.

In recent years, techniques from machine learning have been used to solve some of the analytically/numerically complex problems in quantum foundations. In particular, the methods from reinforcement learning and supervised learning have been used for determination of the maximum quantum violation of various Bell inequalities, the classification of experimental statistics in local/nonlocal sets, training AI for playing Bell nonlocal games, using hidden neurons as hid- den variables for completion of quantum theory, machine learning-assisted state classification, and so forth.

The characterization of the local set for the convex scenario via Bell inequalities becomes intractable as the complexity of the underlying scenario grows (in terms of the number of parties, measurement settings and outcomes). For networks where several independent sources are shared among many parties, the situation gets increasingly worse. The local set is remarkably non-convex, and hence proper analytical and numerical characterization, in general, is lacking. Applying machine learning technique to tackle these issues were studied by Canabarro et al. and Krivachy et al. In the work by Canabarro et al., the detection and characterization of nonlocality is done through an ensemble of multilayer perceptrons blended with genetic algorithms.

In these scenarios, it has been shown[4] 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. We will be using this method of Locally Hidden Variable Neural Network Models for our study.

The existing research had focused on using Bell states as the quantum resource along with joint entangled measurements, we aim to generalize this to X states. We aim to find the best measurement settings and whether we can uncover some global symmetry or property from our study.

Methodology / Review of Studied Distributions

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[2] 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.

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).

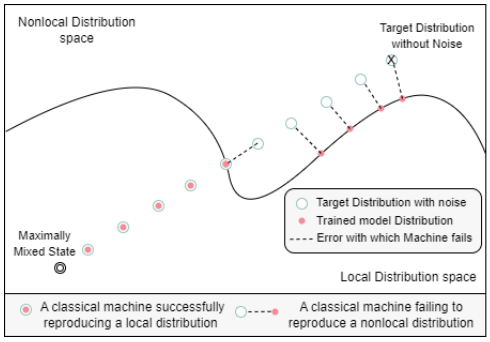


Figure 5: Local to Nonlocal transition through noise visibility

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 [4], 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.

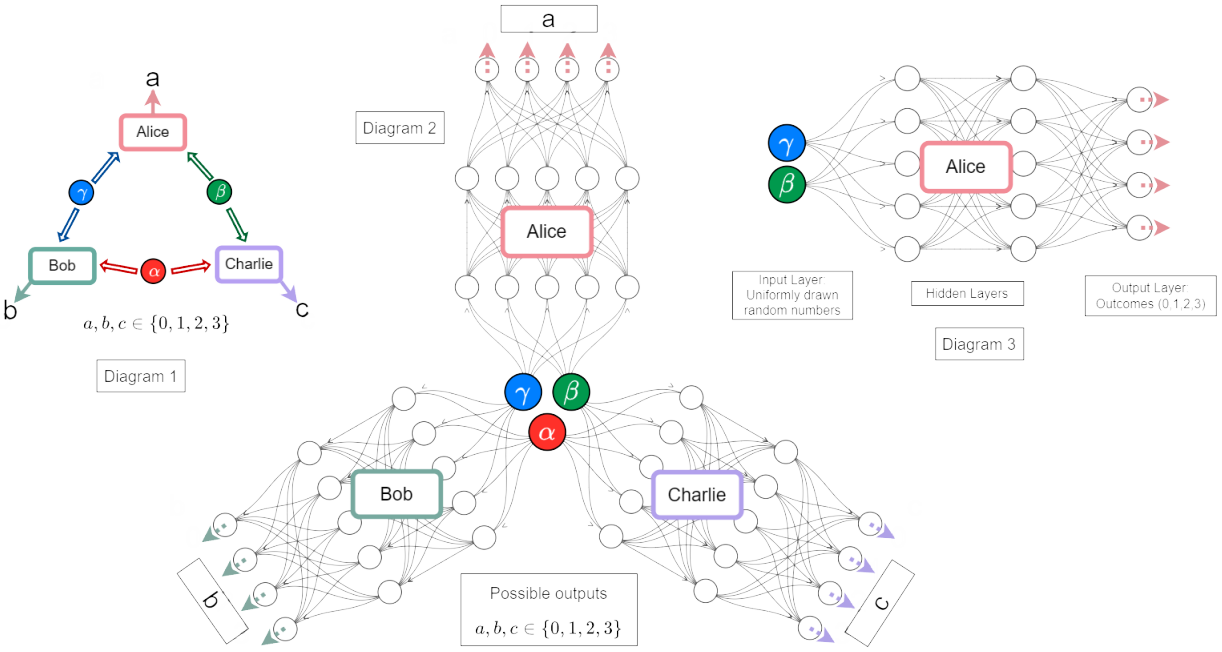


Figure 6: Local Hidden Variable Neural Network

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 probabilities $P_A(a|n_2, n_3), P_B(b|n_1, n_3)$ 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 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.

Fritz Distribution

We use the Fritz distribution[6] 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.

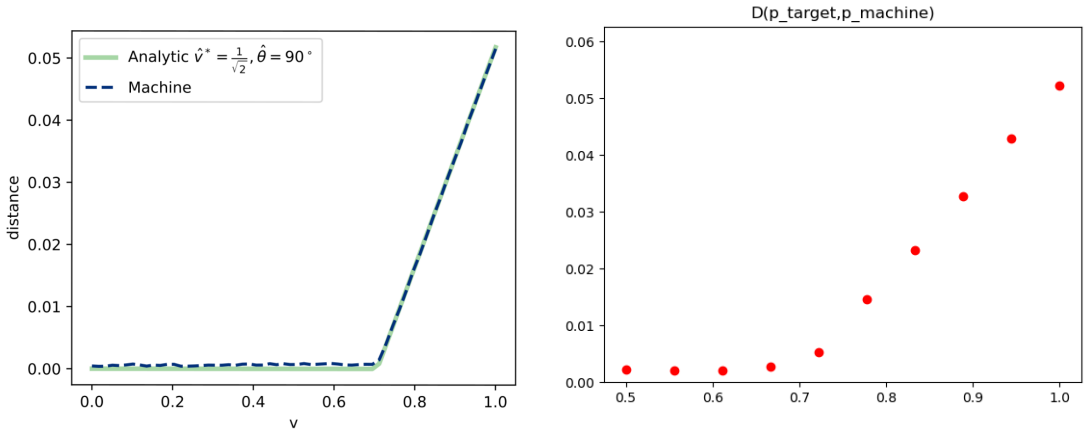
For the noise model, we can introduce finite visibility for the singlet shared between A and B, using the Werner state; where v is the visibility

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

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

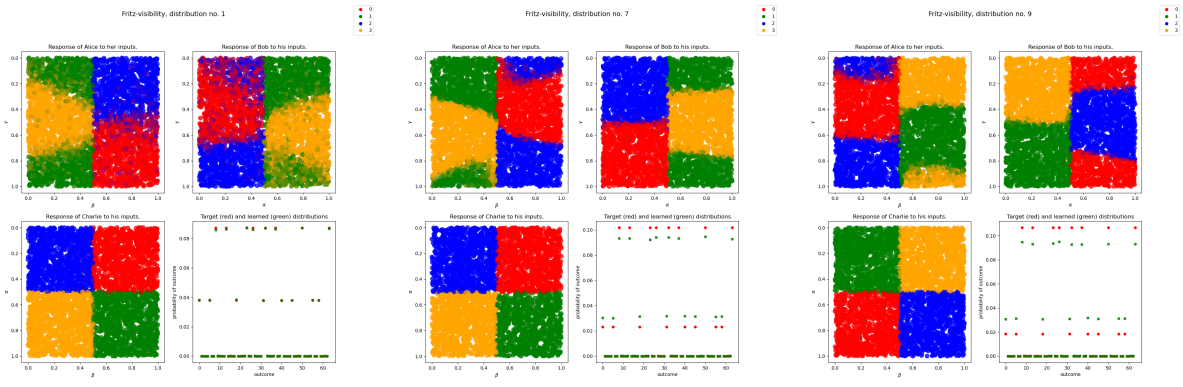
Response functions don't change much after v^*.This is because the machine finds same distribution for the nonlocal distributions that lie outside the local set.

This is due to the special case of Bell nonlocality and Fritz[4] where the local set is a polytope. Figure 5 shows how it will behave.(6)



(a) [4] (b)

Figure 7: 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 $\theta = 90°$

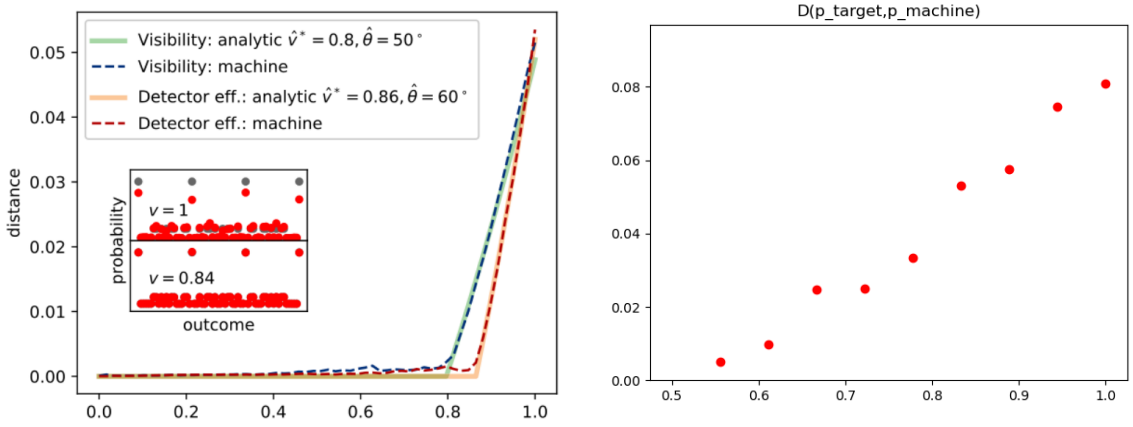


(a) v=0.1 (b) v=0.8 (c) v=1.0

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

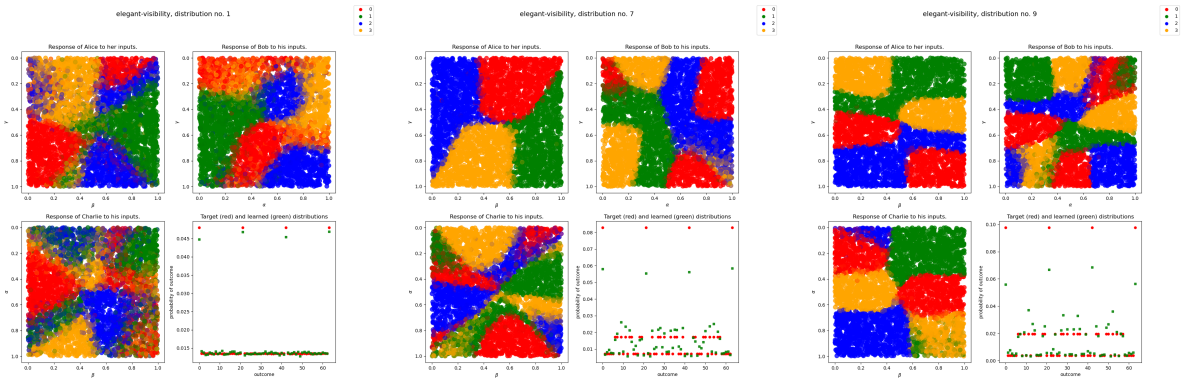
Elegant Distribution

This distribution[7] 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) [4] (b)

Figure 9: 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

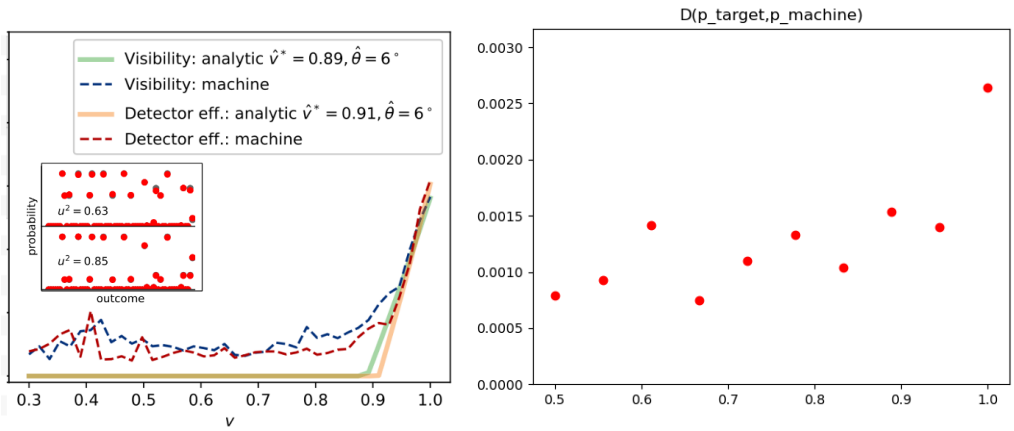
Figure 10: Visualization of response functions and distance between p_t & p_m for different noise parameters

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[4] 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) [4] (b)

Figure 11: Plot of euclidean distance perceived by the machine $d_M(v)$ and the analytic distance $\hat{d}$(v)

The Problem

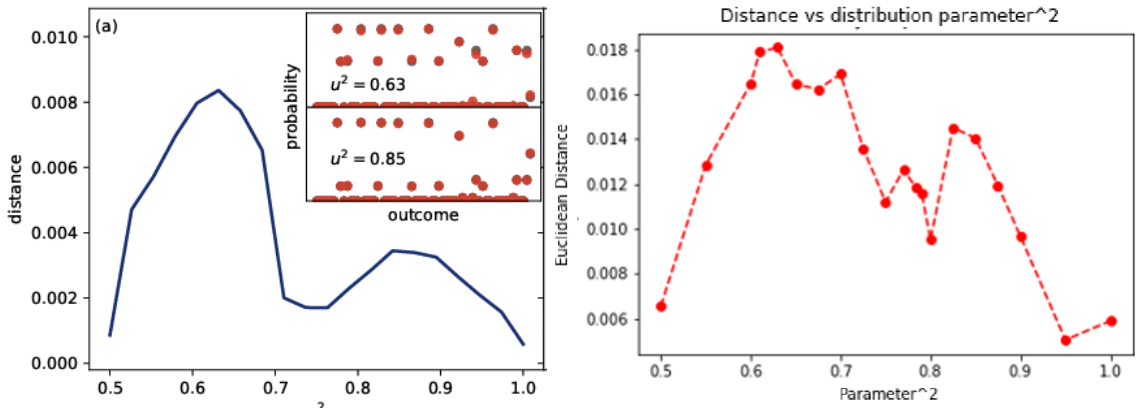
Classifying Nonlocal distributions in Bell and Werner States using joint entanglement measurements

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 [2]:

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

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 [2] 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 [4], 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) [4] (b)

Figure 12: Plot of the distance perceived by the machine, $d_M(v)$ for different values of the entanglement parameter)

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.

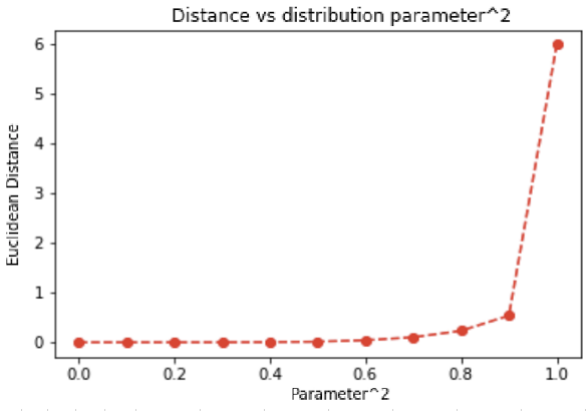


Figure 13: Noisy Werner state

The Problem: 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 entangled 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 parameters. We have also included two sets of entangled measurements instead of just one in the measurement statistics to account for the correlations in the quantum state.

The X states has a total of 7 parameters

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

To understand the X states better and find the best measuremnt setting for X states and possibly the 64 dimensional space. Lets start with a smaller set with a single parameter. Having established our resuls we will move on to the case with three parameters and show that the results are general for a wide range of cases.

These are the quauntum resources we are working with, the state we have studied is a type of X state with two parameters and we have included two parameter joint entanglement measurements.

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

The Measurement basis are

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

The way forward

Phase Zero - Best measurement settings

From the earlier research of Renou et al and T.Krivachy et al we have insight on the area of nonlocality in the triangle network and approaching the systems using LHV neural network models also gave measurement operators wherein we can get these nonlocal correlations. The existing work[4] had focused on using the Bell maximally entangled states with single parameter entangled measurement operators, we instead are exploring the nature of nonlocality with X states and how it traverses this domain with two parameter entangled measurements.

Let us construct explicitly our quantum distribution $P_Q(a, b, c)$ for training our LHV Neural Networks. Each source produces the same pure maximally entangled state of two qubits,

$$|\psi_\gamma\rangle_{A_\gamma B_\gamma} = |\psi_\alpha\rangle_{B_\alpha C_\alpha} = |\psi_\beta\rangle_{C_\beta A_\beta} = \begin{bmatrix} p & 0 & 0 & q \\ 0 & r & s & 0 \\ 0 & s & r & 0 \\ q & 0 & 0 & p \end{bmatrix}$$

Note that each party recieves two independent qubit subsystems; for instance Alice recieves subsystems A_β and A_γ. Next, each party performs a projective quantum measurement in the same basis. In the following, we use the basis (a set depending on our real parameters u & w) given by $|M_{a,b,c}\rangle$. These are the POVM's we are using, and each correspond to an output giving 4 possible output for each measurement.

$$u|00\rangle + (\sqrt{1-u^2})|11\rangle, \quad w|01\rangle + (\sqrt{1-w^2})|10\rangle,$$
$$(\sqrt{1-u^2})|00\rangle - u|11\rangle, \quad (\sqrt{1-w^2})|01\rangle - w|10\rangle$$

with $u^2, w^2 \in [0.5, 1]$. All the four states in the basis are entangled. The global quantum state ψ_g of the triangle network can be found by taking tensor product of the individual sources and

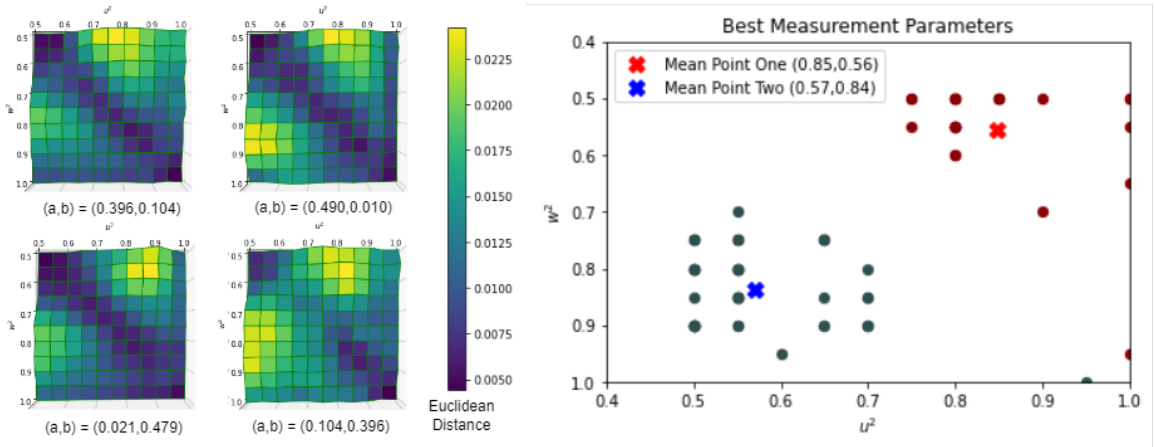


Figure 14: Bell best measurements and points cluster

doing the corresponding network transformation. The statistics of the experiment are given by

$$P_Q(a,b,c) = \langle M_a|\langle M_b|\langle M_c|\psi_g|M_a^t\rangle|M_b^t\rangle|M_c^t\rangle$$

We now use this generated distribution as the target distribution for our LHV Neural Network. Since we have four choices for each of the three agents, we have the probability distribution $P(a,b,c)$ with 64 elements.

Exception

With our initial study and testings, we stumbled upon an unexpected exception when we explore the X states distribution domain. We found that using the LHV neural network model we identified Classically correlated states such as

$$|\psi\rangle = \frac{1}{2}(|00\rangle\langle 00| + |11\rangle\langle 11|)$$

when coupled with the entangled measurements showed quantum nonlocal behaviour in the triangle network.

This means that these distributions couldn't be reproduced by the LHV Neural network and it could be that since entnagled measurements are also a resource in the network scenario, would alone necessitate nonlocality.

We isolated the states that are responsible for this as.

$$|\psi_\gamma\rangle_{A_\gamma B_\gamma} = |\psi_\alpha\rangle_{B_\alpha C_\alpha} = |\psi_\beta\rangle_{C_\beta A_\beta} = \begin{bmatrix} p & 0 & 0 & 0 \\ 0 & r & 0 & 0 \\ 0 & 0 & r & 0 \\ 0 & 0 & 0 & p \end{bmatrix}$$

and the measurement POVM's to be joint Bell entangled measurements

$$u|00\rangle + (\sqrt{1-u^2})|11\rangle, \quad w|01\rangle + (\sqrt{1-w^2})|10\rangle,$$
$$(\sqrt{1-u^2})|00\rangle - u|11\rangle, \quad (\sqrt{1-w^2})|01\rangle - w|10\rangle$$

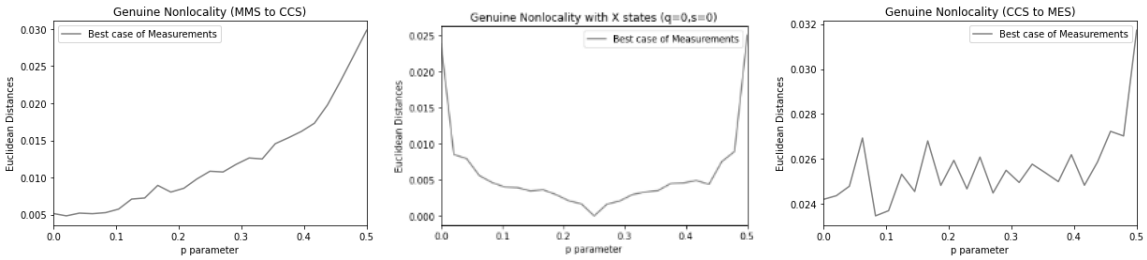


Figure 15: Euclidean distances from Maximally mixed to Maximally Entangled states

We found the Euclidean distances to peak around the states with the parameter $p \in 0.5, 0$. This includes the states

$$|\psi\rangle = \frac{1}{2}(|00\rangle\langle 00| + |11\rangle\langle 11|), \quad \frac{1}{2}(|01\rangle\langle 01| + |10\rangle\langle 10|)$$

This led us to speculate that these state might be an interesting place to explore this conundrum. Focusing on the same we also studied how the euclidean distances vary when we move from maximally mixed states to our classically correlated state, and how it varies from there to maximally entangled states Fig. **??**.

This result isn't intuitively impossible, because in quantum networks entangled measurements are also a quantum resource that plays a role in the nonlocal behavior of the network system. Our problem also aims to prove the extend of this property, from the self testing work on bilocality network it was shown that bilocality network can't express nonlocality with all the sources being separable. We aim to prove if it is actually possible when it comes to a triangle network scenario.

Analytical Disagreement

After, confirming the results of the existing LHV neural network model we have we decided to approach the problem analytically using the $\frac{1}{2}(|00\rangle\langle 00| + |11\rangle\langle 11|)$.

From the Triangle Quantum Network we can send from each sources different qubits to the observers, let us take α as $\rho_{B_1C_2}$, β as $\rho_{A_4C_3}$, and γ as $\rho_{B_6A_5}$. Now by taking the tensor product of each quantum state, we get the global quantum state as $\rho_{B_1A_2} \otimes \rho_{A_4C_3} \otimes \rho_{B_6A_5}$. Now each our participants Alice, Bob, and Charlie does each of their measurements to their shared quantum state with POVM's $P_{AA}^{Alice}, P_{BB}^{Bob}, and P_{CC}^{Charlie}$. Now after adjusting the basis before each of the quantum measurements we get the global quantum state $\rho_{A_5A_4B_1B_6C_3C_2}$.

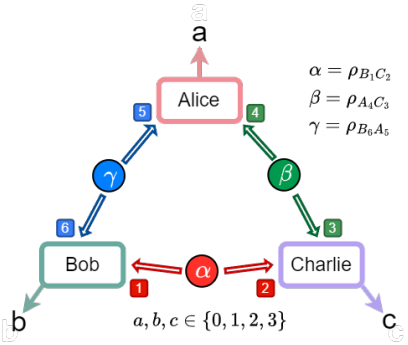


Figure 16: The triangle network

A common test & trivial for checking for nonclassicality is to to take the individual partial traces and see if their tensor product gives back the original global quantum state. But for our global quantum state which is classically correlated and diagonal all the partial traces were maximally mixed states, $|\psi\rangle = \frac{1}{4}(|00\rangle\langle 00| + |11\rangle\langle 11| + |01\rangle\langle 01| + |10\rangle\langle 10|)$, so that is ruled out.

So the global quantum state has to be represented as weighted sum of the quantum states of the individual subsystems, as it is done when the sources are separable states.

$$\rho_{AABBCC} = \sum_{k=1}^{n} p_k \; \rho_{AA}^k \; \rho_{BB}^k \; \rho_{CC}^k$$

$$
\begin{aligned}
p(abc) &= Tr((P^a_{AA} \otimes P^b_{BB} \otimes P^c_{CC})\rho_{AABBCC}) \\
&= Tr((P^a_{AA} \otimes P^b_{BB} \otimes P^c_{CC})(\sum_{k=1}^{n} p_k\ \rho^k_{AA}\ \rho^k_{BB}\ \rho^k_{CC})) \\
&= Tr(\sum_{k=1}^{n} p_k\ P_{AA}\rho^k_{AA} \otimes P_{BB}\rho^k_{BB} \otimes P_{CC}\rho^k_{CC}) \\
&= \sum_{k=1}^{n} p_k\ Tr(P_{AA}\rho^k_{AA} \otimes P_{BB}\rho^k_{BB} \otimes P_{CC}\rho^k_{CC}) \\
&= \sum_{k=1}^{n} p_k\ Tr(P_{AA}\rho^k_{AA})\ Tr(P_{BB}\rho^k_{BB})\ Tr(P_{CC}\rho^k_{CC})
\end{aligned}
$$

Now, let us consider our case with the classically correlated state and its global quantum state.

$$|\psi\rangle = \frac{1}{2}(|00\rangle\langle 00| + |11\rangle\langle 11|)$$

$$
\begin{aligned}
\rho_{AABBCC} &= \frac{1}{8}|00\rangle\langle 00| \otimes |00\rangle\langle 00| \otimes |00\rangle\langle 00| + \frac{1}{8}|00\rangle\langle 00| \otimes |01\rangle\langle 01| \otimes |10\rangle\langle 10| + \\
&= \frac{1}{8}|01\rangle\langle 01| \otimes |01\rangle\langle 01| \otimes |11\rangle\langle 11| + \frac{1}{8}|01\rangle\langle 01| \otimes |00\rangle\langle 00| \otimes |01\rangle\langle 01| + \\
&= \frac{1}{8}|10\rangle\langle 10| \otimes |10\rangle\langle 10| \otimes |00\rangle\langle 00| + \frac{1}{8}|11\rangle\langle 11| \otimes |10\rangle\langle 10| \otimes |01\rangle\langle 01| + \\
&= \frac{1}{8}|10\rangle\langle 10| \otimes |11\rangle\langle 11| \otimes |10\rangle\langle 10| + \frac{1}{8}|11\rangle\langle 11| \otimes |11\rangle\langle 11| \otimes |11\rangle\langle 11|
\end{aligned}
$$

Now, using this quantum state in the equation, we get

$$
\begin{aligned}
p(abc) &= \sum_{k=1}^{n} Tr(P_{AA}\rho^k_{AA})\ Tr(P_{BB}\rho^k_{BB})\ Tr(P_{CC}\rho^k_{CC}) \\
&= \frac{1}{8}Tr(|00\rangle\langle 00|P_{AA})\ Tr(|00\rangle\langle 00|P_{BB})\ Tr(|00\rangle\langle 00|P_{CC}) + \\
&= \frac{1}{8}Tr(|00\rangle\langle 00|P_{AA})\ Tr(|01\rangle\langle 01|P_{BB})\ Tr(|10\rangle\langle 10|P_{CC}) + \\
&= \frac{1}{8}Tr(|01\rangle\langle 01|P_{AA})\ Tr(|01\rangle\langle 01|P_{BB})\ Tr(|11\rangle\langle 11|P_{CC}) + \\
&= \frac{1}{8}Tr(|01\rangle\langle 01|P_{AA})\ Tr(|00\rangle\langle 00|P_{BB})\ Tr(|01\rangle\langle 01|P_{CC}) + \\
&= \frac{1}{8}Tr(|10\rangle\langle 10|P_{AA})\ Tr(|10\rangle\langle 10|P_{BB})\ Tr(|00\rangle\langle 00|P_{CC}) + \\
&= \frac{1}{8}Tr(|11\rangle\langle 11|P_{AA})\ Tr(|10\rangle\langle 10|P_{BB})\ Tr(|01\rangle\langle 01|P_{CC}) + \\
&= \frac{1}{8}Tr(|10\rangle\langle 10|P_{AA})\ Tr(|11\rangle\langle 11|P_{BB})\ Tr(|10\rangle\langle 10|P_{CC}) + \\
&= \frac{1}{8}Tr(|11\rangle\langle 11|P_{AA})\ Tr(|11\rangle\langle 11|P_{BB})\ Tr(|11\rangle\langle 11|P_{CC})
\end{aligned}
$$

Since for the state $|\psi\rangle = \frac{1}{2}(|00\rangle\langle 00| + |11\rangle\langle 11|)$ each participant (Alice, Bob and Charlie) have only two choices we can assign those choices as 0 & 1 with equal probability.

$$p(\alpha = 0) = p(\alpha = 1) = \frac{1}{2}$$
$$p(\beta = 0) = p(\beta = 1) = \frac{1}{2}$$
$$p(\gamma = 0) = p(\gamma = 1) = \frac{1}{2}$$

Now for any quantum multipartite network system, if the joint probability distribution follows the traingle network locality constraint, then there exists a local hidden variable model for the same.

$$p(a, b, c) = \int d\alpha \int d\beta \int d\gamma \; p_A(a|\beta, \gamma) \; p_B(b|\gamma, \alpha) \; p_C(c|\alpha, \beta)$$

Let us use the configuration we assigned with the local hidden varibale model equation.

$$p(abc) = \sum_{\alpha,\beta,\gamma=0,1} p(\alpha)p(\beta)p(\gamma) \; p(a|\beta\gamma) \; p(b|\gamma\alpha \; p(c|\alpha\beta$$

$$\begin{aligned}
&p(a|\beta = 0, \gamma = 0) = Tr(|00\rangle\langle 00|P_A), \quad p(a|\beta = 0, \gamma = 1) = Tr(|10\rangle\langle 10|P_A),\\
&p(a|\beta = 1, \gamma = 0) = Tr(|01\rangle\langle 01|P_A), \quad p(a|\beta = 1, \gamma = 1) = Tr(|11\rangle\langle 11|P_A),\\
&p(b|\gamma = 0, \alpha = 0) = Tr(|00\rangle\langle 00|P_B), \quad p(b|\gamma = 0, \alpha = 1) = Tr(|01\rangle\langle 01|P_B),\\
&p(b|\gamma = 1, \alpha = 0) = Tr(|10\rangle\langle 10|P_B), \quad p(b|\gamma = 1, \alpha = 1) = Tr(|11\rangle\langle 11|P_B),\\
&p(c|\alpha = 0, \beta = 0) = Tr(|00\rangle\langle 00|P_C), \quad p(c|\alpha = 0, \beta = 1) = Tr(|01\rangle\langle 01|P_C),\\
&p(c|\alpha = 1, \beta = 0) = Tr(|10\rangle\langle 10|P_C), \quad p(c|\alpha = 1, \beta = 1) = Tr(|11\rangle\langle 11|P_C)
\end{aligned}$$

By putting these values into the hidden variable equation, we get the $p(abc)$ joint probability distribution for our classically correlated state, which means that there is indeed a hidden variable model for it.

This is in stark contradiction to what our LHV neural network model found. Following this we compiled the same distribution using bigger training sets and fine tuned our results, but the LHV NN model failed to achieve the degree of euclidean distances for the classically correlated state like it achieved for the maximally mixed state.

New LHV Neural Network from Analytical Proof

Looking back into the constraints of the traingle lhv neural network model, those are the partially connected nature of the nerual network imposing the networks input constraints. Alongwith the inputs being from a random uniform distribution (mimicking hidden variables) and the outputs following the constraints of a probability distribution. These together enforce the constraints responsible for the neural network not being to learn nonlocal distributions.

We believe that the primary reason the model failed to categorise the earlier classical state distribution as local was that the neural network might be having stronger constraints that

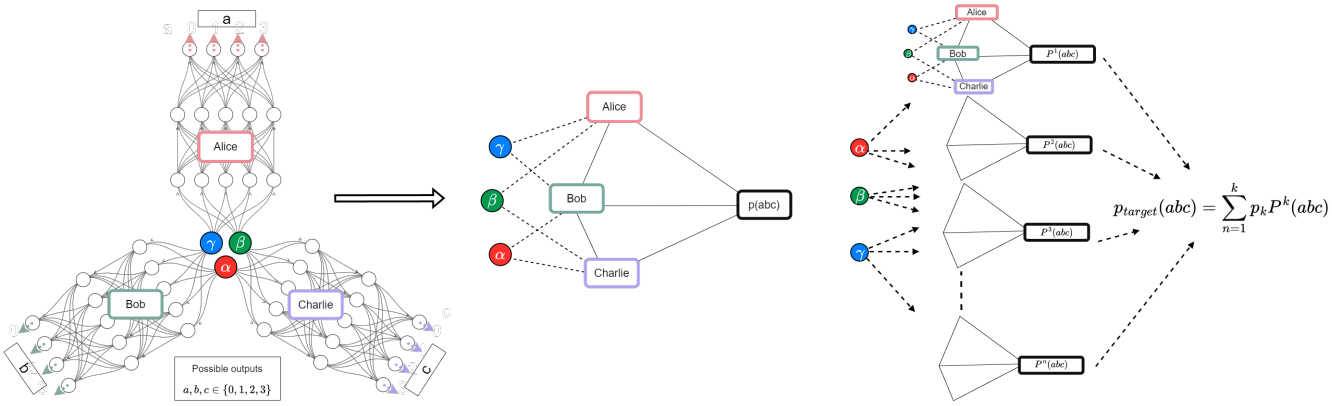


Figure 17: New Traingle LHV Neural Network model

not only limit the model's capacity to learn nonclassical distributions but also certain classical ones. From our results we can see that the model faces no difficulty in distinuishing maximally mixed states and maximally entnagled states.

So we decided to consider the general case of separable states and explore. We know that if the state is separable we can express the same as

$$\rho_{AABBCC} = \sum_{k=1}^{n} p_k \ \rho_{AA}^k \ \rho_{BB}^k \ \rho_{CC}^k$$

$$\begin{aligned}
p(abc) &= Tr((P_{AA}^a \otimes P_{BB}^b \otimes P_{CC}^c)\rho_{AABBCC}) \\
&= Tr((P_{AA}^a \otimes P_{BB}^b \otimes P_{CC}^c)(\sum_{k=1}^{n} p_k \ \rho_{AA}^k \ \rho_{BB}^k \ \rho_{CC}^k)) \\
&= Tr(\sum_{k=1}^{n} p_k \ P_{AA}\rho_{AA}^k \otimes P_{BB}\rho_{BB}^k \otimes P_{CC}\rho_{CC}^k) \\
&= \sum_{k=1}^{n} p_k \ Tr(P_{AA}\rho_{AA}^k \otimes P_{BB}\rho_{BB}^k \otimes P_{CC}\rho_{CC}^k) \\
&= \sum_{k=1}^{n} p_k \ Tr(P_{AA}\rho_{AA}^k) \ Tr(P_{BB}\rho_{BB}^k) \ Tr(P_{CC}\rho_{CC}^k)
\end{aligned}$$

As we can see now that the probability distribution is a result of the sum of several other probability distributions. The previous model in this manner best works with states that are pure states, and hence does not include this particular freedom in learning the distributions when considering X states. Since the model tried to learn a single distribution, whereas the target is a sum of many such distributions, it failed.

So in this manner we used this analytical setting to create a new LHV machine learning model capable of distinguishing nonlocality in not just pure states but also mixed ones.

This new model consists of k sets of neural networks that can distinguish nonlocality in any global quantum state of rank k density matrix and local measurements(including joint entnagled measurements)

So, like in the figure 17 we can use two models in unison and distinguish genuine network nonlocality usin a global quantum state with rank 2 and local measurements. Looking back into our earlier classically correlated state we had the target distribution split as a sum of 8 other distributions of pure states. That is the target distributions state is a mixed state of rank 8, so we will need 8 of these models acting in unison. We will need a maximum of 64 of these parallel models for the entire 64x64 global density matrix of the triangle network.

Results

Phase I - Best measurement settings for one parameter X States

From the earlier research of Renou et al and T.Krivachy et al we have insight on the area of nonlocality in the triangle network and approaching the systems using LHV neural network models also gave measurement operators wherein we can get these nonlocal correlations. The existing work[4] had focused on using the Bell maximally entangled states with single parameter entangled measurement operators, we instead are exploring the nature of nonlocality with X states and how it traverses this domain with two parameter entangled measurements. Let us construct explicitly our quantum distribution $P_Q(a, b, c)$ for training our LHV Neural Networks. Each source produces the same pure maximally entangled state of two qubits,

$$|\psi_\gamma\rangle_{A_\gamma B_\gamma} = |\psi_\alpha\rangle_{B_\alpha C_\alpha} = |\psi_\beta\rangle_{C_\beta A_\beta} = \begin{bmatrix} a & 0 & 0 & a \\ 0 & b & b & 0 \\ 0 & b & b & 0 \\ a & 0 & 0 & a \end{bmatrix}$$

Unitary Trace condition of Density Matrices, $Tr(\psi_{\alpha,\beta,\gamma}) = 1$. So $b = 0.5 - a$.

We can get this representation by considering a linear combination of two of the Bell states. (Bell diagonal states)

$$|\phi^+\rangle = \frac{|00\rangle + |11\rangle}{\sqrt{2}}, \qquad |\psi^+\rangle = \frac{|01\rangle + |10\rangle}{\sqrt{2}},$$

We have gradually increased the value of a from 0 to 0.5; $b = 0.5 - a$ because of the trace constraint, so effectively we are dealing with one parameter in this case of X states. While we vary a 0.0 to 0.25 and 0.25 to 0.5 is the same with just the a and b values swapped. The states with a = 0 & 0.5 are Bell states and a = 0.25 is a purely local state. Note that each party recieves two independent qubit subsystems; for instance Alice recieves subsystems A_β and A_γ. Next, each party performs a projective quantum measurement in the same basis. In the following, we use the basis (a set depending on our real parameters u & w) given by $|M_{a,b,c}\rangle$. These are the POVM's we are using, and each correspond to an output giving 4 possible output for each measurement.

$$\begin{aligned} &u|00\rangle + (\sqrt{1-u^2})|11\rangle, \quad w|01\rangle + (\sqrt{1-w^2})|10\rangle, \\ &(\sqrt{1-u^2})|00\rangle - u|11\rangle, \quad (\sqrt{1-w^2})|01\rangle - w|10\rangle \end{aligned}$$

with $u^2, w^2 \in [0.5, 1]$. All the four states in the basis are entangled. The global quantum state ψ_g of the triangle network can be found by taking tensor product of the individual sources and doing the corresponding network transformation. The statistics of the experiment are given by

$$P_Q(a, b, c) = \langle M_a|\langle M_b|\langle M_c|\psi_g|M_a^t\rangle|M_b^t\rangle|M_c^t\rangle$$

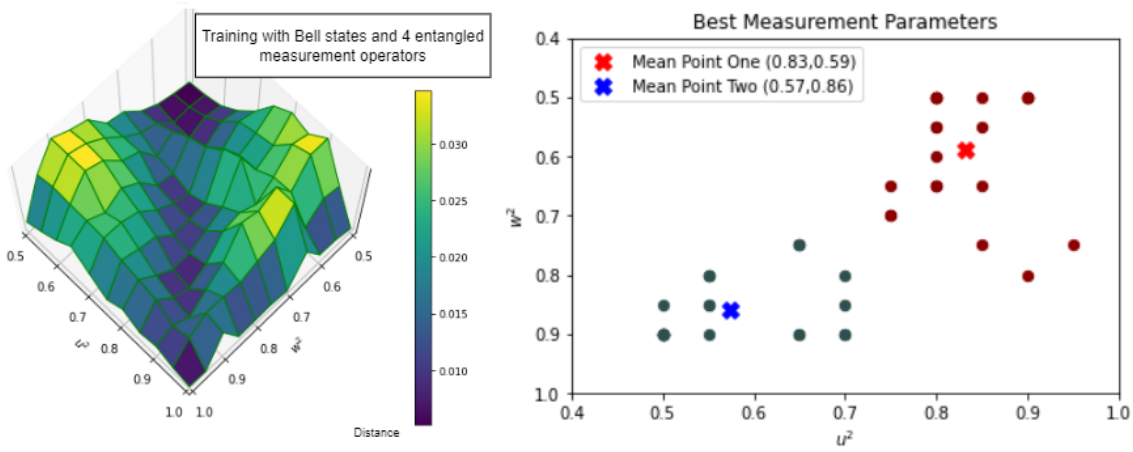


Figure 18: X states best measurements and points cluster

We now use this generated distribution as the target distribution for our LHV Neural Network. Since we have four choices for each of the three agents, we have the probability distribution $P(a, b, c)$ with 64 elements.

We have three main results

- Best measurement settings
- The two best choices of the distribution parameters have symmetry.
- Noise robustness of the quantum states.

Firstly by parameterizing the quantum states and calculating the quantum distribution using the joint measurements, we use these values to train our LHV Neural Network using the classical random values as the sources. Given a target distribution p_t, the neural network provides an explicit model for a distribution p_m, which according to machine is the closest local distribution to p_t. p_m is guaranteed to be from local set by construction. The neural network will almost never exactly represent the p_t distribution exactly. This is because p_m is learned by sampling a distribution a finite no of times, and the learning technique does not guarantee conversion to a global optimum. We repeat the process for a range of the parameterized values, and by doing this and recording the trained distributions, we can then find the euclidean distance between each trained and target distribution.

Since in here the ability to reproduce a distribution is the same as deeming the distribution classical, the euclidean distance between the target and trained distribution acts as a measure of nonlocality. Making use of this, we can then find and study the topology of the surface to identify the maximum genuine nonlocality possible for this configuration. It is especially more beneficial to use some behavior statistics like the euclidean distance between the distributions to study the behavior of the same. And this transition in the machine's behavior when we transition between local and nonlocal distributions are especially robust and informative.

By using the euclidean distance as a measure of nonlocality, we found the best measurement setting to be (18)

$$(u^2, w^2) = (0.555, 0.830) \ \& \ (0.820, 0.540)$$

of which the former measurement have the relatively best expression in all of X states, which we will revisit next. Here in Figure 18, the first set of graphs shows the euclidean distance of X states (single parameter) for varying parameters of u & w, in the second we have actually complied and found the peak euclidean distance values for varying a & b values, and found the centroid. This value $(u^2, w^2) = (0.555, 0.830)$ & $(0.820, 0.540)$ is the best mesaurement setting for our sample space. We tested the same method on the Bell States, and the third graph shows the same two peak significant peaks of euclidean values and the centroid of the peak values in the fourth graph, which is approx the same as before (which is expected as Bell states are also X states).

In 19, I have used the two best measurement choices to measure the quantum state while varying its single parameter across X states, and we can see that it approx aligns with the surface of local minimas. There is a small error factor due to the training nature of neural network.

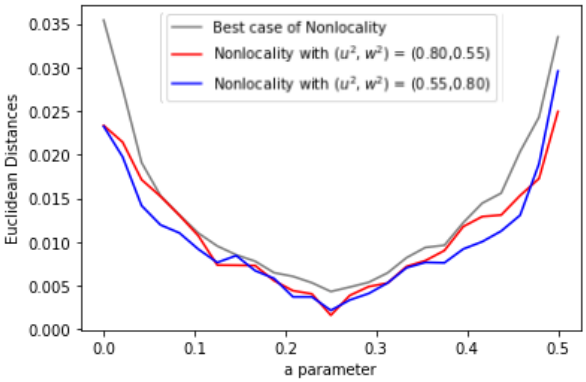


Figure 19: The two best choice of measurements

Also, you can see the best measurement coordinates of the case shown in 18(graph 1) are in fact swaps of each other. We believe this in fact due to the case of b parameter being related to a. Also if we look into the analytical expression of the probability distribution, we can find that

$$P(a, b, u, w) = P(b, a, w, u)$$

And since from here, we get $P(u, w) = P(w, u)$ it gives us $P(a, b) = P(b, a)$ which is in fact true and can be observed when we vary a from 0.0 to 0.25 & 0.25 to 0.5; where these distributions swap.

Noise Robustness study of Nonlocal Quantum Distributions

We now add noise using the Werner states, and explore the noise robustness of these states.

$$\rho(v) = v|\psi^-\rangle\langle\psi^-| + (1 - v)I/4 \tag{9}$$

For the noise model, we can introduce finite visibility for the quantum state shared between each two parties, using the Werner state; where v is the visibility. And, $I/4$ denotes the maximally mixed state of two qubits. By taking the quantum distribution believed to be nonlocal & adding some degree of noisyness controlled by the parameter v with $p_t(v = 0)$ being the completely noisy local distribution and $p_t(v = 1)$ being the noiseless most nonlocal distribution.

By adding noise, we can guarantee, at some point it enters the local set v ¿ v*

$$d(p_t, p_m) = \sqrt{\sum_{a,b,c} [p_t(abc) - p_m(abc)]^2}$$

like in figure 20. We can see that although the target distribution without noise lies outside the local domain, by subsequently adding noise we can bring this distribution to the local domain. And an important fact to consider is the terrain of the local-nonlocal boundary can't be assumed like in Standard Bell scenaros where we have a polytope. A clear liftoff of the distribution $d_M(v) = d(p_t(v), p_m(v))$ at some point is a signal that we are leaving the local set. $d_m(v)$ distance - we can deduce a lot.

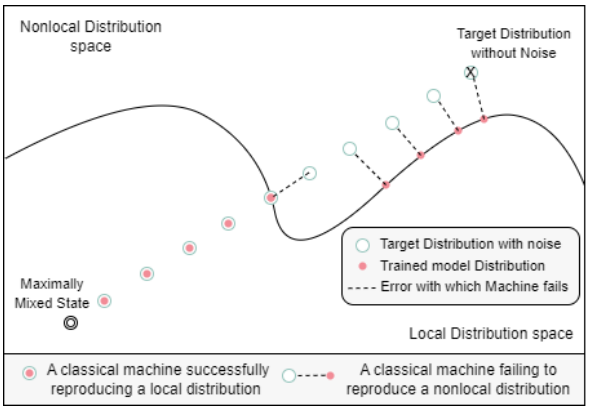


Figure 20: Adding noise to nonlocal distribution

Also when the quantum distribution lies in the local domain in 20, it means that this particular distibution is learnable and can be reproduced by a classical machine learning model. The nonlocal distribution although unlearnable casts a local distribution when the machine tries to learn it, this unsurprisingly trains to be the closest distribution to the nonlocal one.

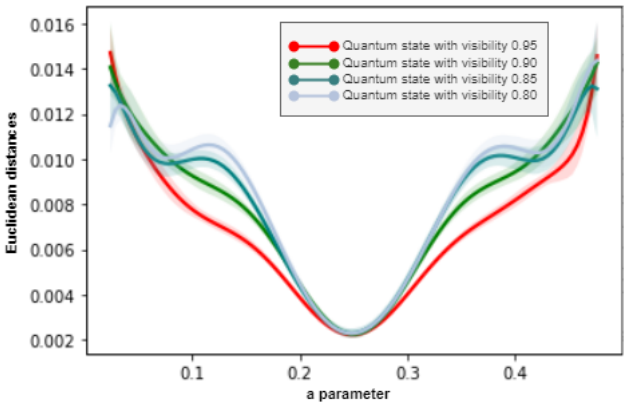


Figure 21: Euclidean distances for varying noise visibilities

Next we also plotted the euclidean distances of quantum states while varying their noise visibility and state parameter a. We can see that in Fig. 21 the euclidean distances is minimal when a = b = 0.25, here the nonlocal nature vanishes and the euclidean distance peaks at the extremes. Adding noise generally decreases the nonlocality here, but we were able to see that for certain ranges of a in Fig. 21, adding noise actually increased the euclidean distances. This behavior seems to increase rapidly in the beginning and later slowly after reaching a threshold of 0.80.

This behavior aligns with a later result in this report; we believe that these states upon being added noise form a quantum state that is more nonlocal in nature, hence show a clear difference in euclidean distances when compared to other states.

Phase II: X States with three parameters

Now we can move on to X states with three parameters, we can prepare this state by taking a linear combination of all four Bell states, which we also call the Bell diagonal states. By taking a linear combination of all

$$|\phi^+\rangle = \frac{|00\rangle + |11\rangle}{\sqrt{2}}, \qquad |\psi^+\rangle = \frac{|01\rangle + |10\rangle}{\sqrt{2}},$$
$$|\phi^-\rangle = \frac{|00\rangle - |11\rangle}{\sqrt{2}}, \qquad |\psi^-\rangle = \frac{|01\rangle - |10\rangle}{\sqrt{2}}$$

We can group the coefficients to give us four parameters. Considering the trace unity condition we have in total have three parameters. Unitary Trace condition of Density Matrices, $Tr(\psi_{\alpha,\beta,\gamma}) = 1$. So $r = 0.5 - p$.

$$|\psi_\gamma\rangle_{A_\gamma B_\gamma} = |\psi_\alpha\rangle_{B_\alpha C_\alpha} = |\psi_\beta\rangle_{C_\beta A_\beta} = \begin{bmatrix} p & 0 & 0 & q \\ 0 & r & s & 0 \\ 0 & s & r & 0 \\ q & 0 & 0 & p \end{bmatrix}$$

We deploy the same set of joint Bell entangled measurement settings as before. These are the POVM's we are using, and each correspond to an output giving 4 possible output for each measurement.

$$u|00\rangle + (\sqrt{1-u^2})|11\rangle, \quad w|01\rangle + (\sqrt{1-w^2})|10\rangle,$$
$$(\sqrt{1-u^2})|00\rangle - u|11\rangle, \quad (\sqrt{1-w^2})|01\rangle - w|10\rangle$$

with $u^2, w^2 \in [0.5, 1.0]$. We follow the same methodology here with the 64 element $P_Q(a, b, c)$ distribution and using it to train the LHV Neural Network. By taking the condition that each eigenstate should be positive to be a valid quantum state, we have the conditions $p \geq q$ & $p \leq 0.5 - s$. We explored the case with different values of p while fixing the other two parameters. We varied the value from 0.0 to 0.5 for the parameters. We were able to identify three main results, in this study.

- First, we found out that the best measurement setting for using the joint entangled measurements and getting nonlocality without inputs is on a case to case basis when considering the X states. A general pattern is not found here, but we can find that the pattern in our previous scenario do carry down here, but the strength of euclidean distances suggest that there is no distinct pattern, as you can see in 22
- Secondly, although the best measurement setting is on a case to case basis, like we mentioned earlier few of the patterns we saw are repeated here, and the best measurement setting $(w^2, u^2) = (0.555, 0.830)$ gives us a good confidence level for attaining nonlocality without inputs.

- Third, we found which of these states have a pronounced expression of nonlocality, and categorised these states to specific quantum states. Along with the best measurement settings for these states, we also found that the measurement setting $(w^2, u^2) = (0.555, 0.830)$ remains the best for these states.

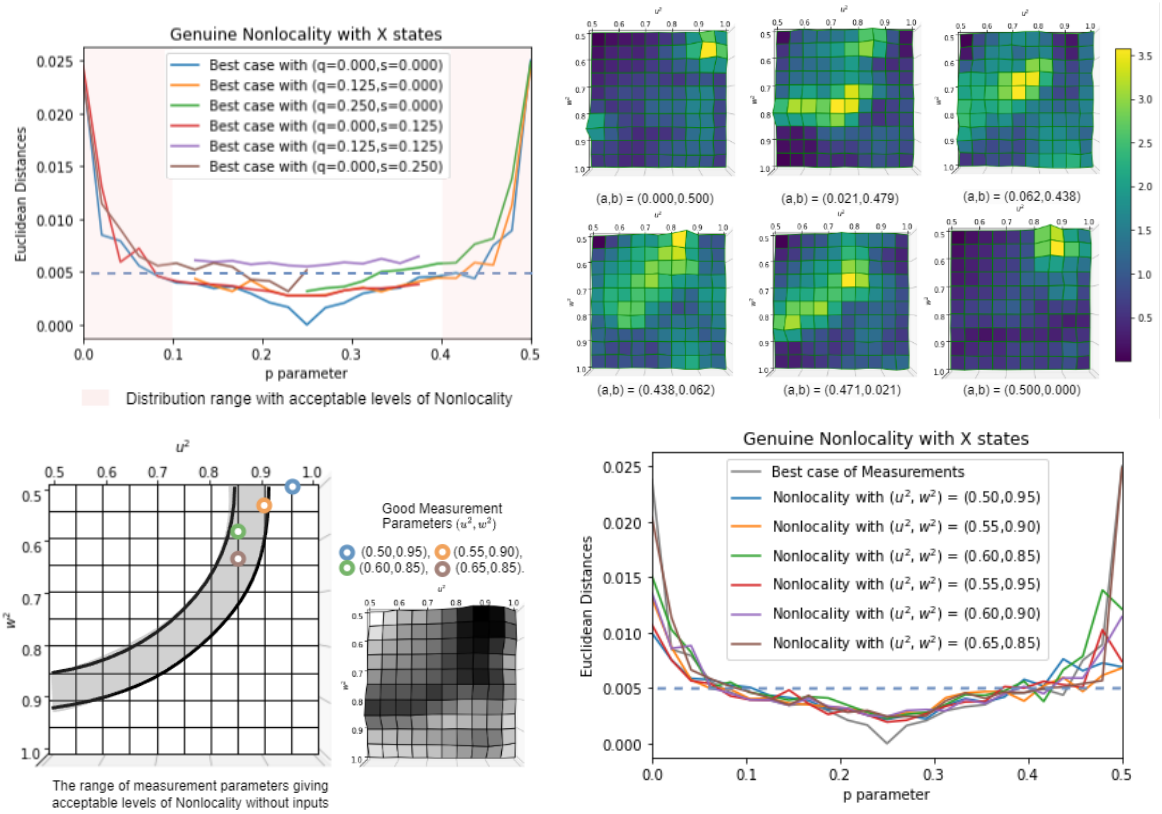


Figure 22: Bell and X states best measurements and points cluster

In Fig. 22 we show the pattern emerging from plotting the best measurement settings for X states with three parameters. In graph 1, we plotted the various X state sets and we have highlighted the quantum distributions showing acceptably high euclidean distances/nonlocal behavior. And from this observation we can see that these states lie in the range of p parameter in $[0.0, 0.1]$ and $[0.4, 0.5]$. This also takes into consideration that $p \geq q$ & $p \leq 0.5 - s$ and we can see that the q values tends to show better nonlocality when it is comparable to p; that is the nonlocal states are attracted to Bell states, which is not surprising.

Like we said before in Fig. 22 the graph 1 shows the quantum states showing acceptably high levels of nonlocality with the parameter $p \in [0.0, 0.1]U[0.4, 0.5]$; where $r = 0.5 - p, r \in [0.5, 0.4]U[0.1, 0.0]$. We found these states to be those with their $q\&s$ parameters zero while following the constraint $q \leq p, s \leq r$.

$$|\psi_\gamma\rangle_{A_\gamma B_\gamma} = |\psi_\alpha\rangle_{B_\alpha C_\alpha} = |\psi_\beta\rangle_{C_\beta A_\beta} = \begin{bmatrix} p & 0 & 0 & q \\ 0 & r & 0 & 0 \\ 0 & 0 & r & 0 \\ q & 0 & 0 & p \end{bmatrix}, \begin{bmatrix} p & 0 & 0 & 0 \\ 0 & r & s & 0 \\ 0 & s & r & 0 \\ 0 & 0 & 0 & p \end{bmatrix}$$

We took these states and found their best measurement setting in graph 2, from which we

can make out a pattern (graph 3). We found these best measurement settings can be sampled from this pattern and on inputting these settings we can find their euclidean distances to match that of the best case scenario (graph 4). We found that of the sample space the best fit for the best measurement setting remained to be the $(u^2, w^2) = (0.850, 0.550)$ case.

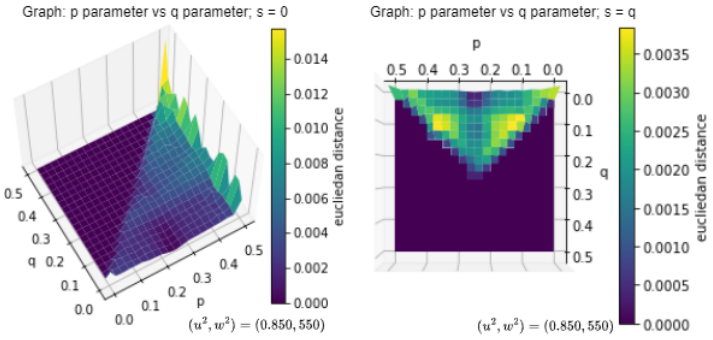
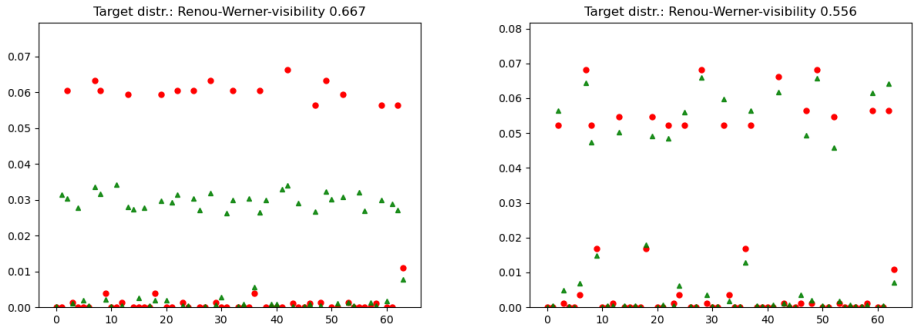


Figure 23: p,q vs euclidean distances

To back our results, we also ran computations with $s = 0$ and $s = q$, while varying the p and q parameters, all while fixing the measurement settings to our best case $(u^2, w^2) = (0.850, 0.550)$.

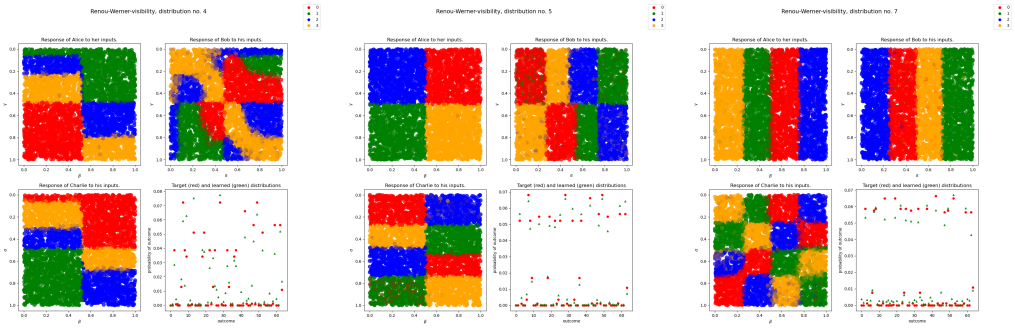
$$|\psi_\gamma\rangle_{A_\gamma B_\gamma} = |\psi_\alpha\rangle_{B_\alpha C_\alpha} = |\psi_\beta\rangle_{C_\beta A_\beta} = \begin{bmatrix} p & 0 & 0 & q \\ 0 & r & 0 & 0 \\ 0 & 0 & r & 0 \\ q & 0 & 0 & p \end{bmatrix}, \begin{bmatrix} p & 0 & 0 & q \\ 0 & r & q & 0 \\ 0 & q & r & 0 \\ q & 0 & 0 & p \end{bmatrix}$$

In Fig. 23 graph 1 we can see that the peak aligns with the Bell state at $p = q = 0.5$ and slowly decreases as q value decreases. In graph 2 the euclidean distances are lowest at $p = 0.25$ which here is the maximally mixed state. The two peaks also agree with our existing observations.



(a) (b)

Figure 24: The distributions show that only a few elements of the 64 probability distributions are responsible for nonlocal behavior



(a) (b) (c)

Figure 25: We also got interesting symmetric behavior in the response functions of these distributions

From the Euclidean distances, we can 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.

Conclusion

We presented the best measurement settings that gives the maximal genuine nonlocality with triangle quantum networks.

Here the Renou et al distribution we present represents a form of quantum nonlocality that is genuine to the network configuration. Here the scenario fundamentally differs from relying on a standard bell inequality. We were able to reproduce the results of Renou et al and T.Krivachy et al by using our improved neural network architecture to classify nonlocal states. For the Renou et al distribution particularly the nonlocal and local distributions are especially close and classifying the states was much more difficult; hence the neural network architecture had to be adjusted with bigger training data sets and extra fine-tuning steps.

This existing work had focused on using the Bell maximally entangled states with single parameter entangled measurement operators, we expanded this to the nature of nonlocality with X states and how it traverses this domain with two parameter entangled measurements. We found that the best measureemnt setting we identified with Bell states also transfers to general case of X states with single parameter. Here we found that for these two sets of quantum states there are two best choices of mesauremnt setting which are symmetric; $(u^2, w^2) = (0.820, 0.540)$ & $(0.555, 0.830)$. We were able to confirm that the valley in between these two peaks where $u^2 = w^2$ gives no nonlocality. We achieved this by parameterizing the quantum states and calculating the quantum distribution using the joint measurements, we used these values to train our LHV Neural Network using the classical random values as the input sources. . Given a target distribution pt, the neural network provides an explicit model for a distribution pm, which according to machine is the closest local distribution to pt. pm is guaranteed to be from local set by construction.

The method we use attacks the problem from a different angle. It relaxes both the discrete hidden variable and deterministic response function assumptions which are made by other methods. The complexity of the problem now boils down to the response function of the observers - each of which is represented by a feedforward neural network. Though an approximate one, one can increase the methods precision by increasing the size of the neural network. Due to universal approximation theorems we are guaranteed to be able to represent essentially any function with arbitrary precision.

We also conducted noise robustness studies by controlling the nonlocality measure by adding Werner noise on the nonlocal distributions; by doing this we got an estimate of the noise visibility required for an X state to show nonlocal behavior. For a single target distribution the machine finds only an upper bound to the distance from the local set. By examining families of target distributions, however, we get a robust signature of nonlocality due to the clear transitions in the distance function, which match very well with the approximately expected distances.

The trilocal set being topologically closed, it is clear that PQ(a, b, c) must have a certain robustness to noise: when adding a sufficiently small amount of local noise to PQ(a, b, c), one should still obtain a quantum distri- bution that is incompatible with any trilocal model. Interestingly, we were also able to see that or certain families of nonlocal distribution adding noise actually increased the incompatibily with any trilocal model upto a threshold.

We expanded our study to three parameter X states, from a linear combination of the four Bell states (Bell diagonal states). Using the two parameter joint entangled Bell state measurements and training the neural network, we found that the best measurement setting for nonlocality without inputs is on a case to case to basis generally when considering the whole of X states. Although a general pattern is not found here, our earlier pattern do carry down here for states that are capable of showing high nonlocality. And the $(w^2, u^2) = (0.555, 0.830)$ comes out as the best case for achieving nonlocality without inputs. Next of the quantum distributions that have a high level of nonlocality we classified those quantum states to a set of groups.

Interestingly we also found that the LHV Neural network model categorised classically correlated states under joint entangled measurements as nonlocal based on its inability to learn said distributions. Following this after rigorously checking the LHV Neural network model with different training sets we confirmed there is a valid difficulty in the training process. We then proceeded to analytically prove that the quantum state in question under joint entangled measurements has a Local hidden variable model. We succeeded in proving that the distibution in question is indeed local and the LHV NN model failed to learn the said distribution. We speculate that there is some nontrivial factor since the problem failed to show improvemnts with diiferent cases despite the natural approximation theorem. Hence this is where we conclude and move onto our future directions for this project.

Future Directions

We are moving forward with the exception we met with when using classically correlated states coupled with joint entangled measurements in the quantum triangle network. As we understand it the inability of the LHV neural network could be better countered using a more biased source set of randum inputs. Currently we are approaching this by training a Generative Network to create our biased network based on the distribution we like to recreate. We are interested in reducing the resources required to reach a good confidence level. Using this, we can explore other network and non-network scenarios. Achieving this will also help us to see if a machine can derive, or at least decipher for a nonlinear Bell-type inequality which is violated by Full Network Nonlocality (FNL) or GNNL.

By improving our approach, we plan to further understand entanglement swapping and GNNL, bilocality network is the place to start to understand the basic fundamentals. In this continued process we plan to reframe the locality constraint of Quantum Networks. Integrating coarse graining tactics to experiment with the output cardinalities of the network. Checking whether the triangle network can give nonlocality without inputs witout all sources with some entnaglement is also an interesting thing to figure out. I like to move forward with this project to other generic networks by incorporating color- matching and token-counting techniques that have been shown to be fruitful in providing insights into the problem.

I am also in the works of introducing a programming sandbox functionality for quantum network nonlocality by consulting and studying compatible approaches capable of simplying the methodolgy and complexity. Having found the optimization in the distribution space especially close to the local-nonlocal boundary quite difficult and since the dimensionality of the terrain is high I have been working on including a hybrid optimizer using metaheuristic algorithms. I am also interested in finding at the end of my current work what kind of input distribution is capable of reproducing the local set of quantum networks.

From my experience of hitting the classically intractable wall of studying quantum systems using classical resources; I am looking forward to understand the Quantum Learning theory behind learning quantum multipartite network systems and emulating such systems. The quantum advantage and the design and development of such algorithms is a really interesting field of study. Bridging classical intractability using quantum-enhanced machine simulation

with quantum learning theory will help us to design better machines capable of answering a whole lot of fundamental questions of our universe. I aspire to combine my interest and expertise in these fields towards my learning goals; to bring together quantum information science, math and machines to discover new facets of our physical reality and our capabilities.

References

[1] Cyril Branciard, Denis Rosset, Nicolas Gisin, and Stefano Pironio. Bilocal versus nonbilocal correlations in entanglement-swapping experiments. *Physical Review A*, 85(3), mar 2012.

[2] Marc-Olivier Renou, Elisa Bäumer, Sadra Boreiri, Nicolas Brunner, Nicolas Gisin, and Salman Beigi. Genuine quantum nonlocality in the triangle network. *Physical Review Letters*, 123(14), sep 2019.

[3]

[4] Tamás Kriváchy, Yu Cai, Daniel Cavalcanti, Arash Tavakoli, Nicolas Gisin, and Nicolas Brunner. A neural network oracle for quantum nonlocality problems in networks. *npj Quantum Information*, 6(1), aug 2020.

[5] Nicolas Brunner, Daniel Cavalcanti, Stefano Pironio, Valerio Scarani, and Stephanie Wehner. Bell nonlocality. *Reviews of Modern Physics*, 86(2):419–478, apr 2014.

[6] Tobias Fritz. Beyond bell's theorem: correlation scenarios. *New Journal of Physics*, 14(10):103001, oct 2012.

[7] Nicolas Gisin. Entanglement 25 years after quantum teleportation: Testing joint measurements in quantum networks. *Entropy*, 21(3):325, mar 2019.

[8] Ryszard Horodecki, Paweł Horodecki, Michał Horodecki, and Karol Horodecki. Quantum entanglement. *Rev. Mod. Phys.*, 81:865–942, Jun 2009.