CERTIFICATION OF STEERABILITY VIA CONVOLUTIONAL NEURAL NETWORKS:

COMPARING REAL, COMPLEX AND QUATERNIONIC NETWORKS FOR A TWO-QUBIT STATE

Arthur C. R. Dutra

Instituto de Matemática, Estatística e Computação Científica, Universidade Estadual de Campinas, Campinas, SP, Brasil

ABSTRACT

In this paper, we explore the certification of steerability for quantum states using neural networks. Quantum steering, a phenomenon where the measurement choices of one party can influence the state of another distant party, is a key feature of quantum theory. Traditionally, dense real neural networks have been employed when training machine learning models on quantum states. However, we aim to investigate the potential benefits of leveraging complex and quaternionic convolutional networks, which naturally align with the mathematical structure of quantum states. By creating the training set with a state-of-the-art steering certification algorithm and incorporating convolutional methods, our approach achieves a balanced accuracy of 94% across the three networks tested, showing that either model can successful when dealing with quantum states. All the code developed and used in this project is available in `https://github.com/turdutra/Steerability-with-NN`

***Index Terms*—** Quantum steering, convolutional neural networks, machine learning

1. INTRODUCTION

In 1935, Einstein, Podolsky, and Rosen published their renowned paper challenging the completeness of quantum theory [1]. They examined a two-particle system in which one observer can measure either the position or momentum of the particle. Due to the entanglement between the particles, measuring the position or momentum of one particle allows the prediction of the corresponding measurement outcome for the second particle, regardless of the distance between them.

This argument sparked enduring debates, and shortly after its publication, Schrödinger identified a fascinating implication. He pointed out that the first observer, by choosing the type of measurement, could influence the state of the second particle, steering it into either an eigenstate of position or momentum [2]. Although this phenomenon cannot transmit information, Schrödinger was still baffled by it.

Today, this phenomenon is known as quantum steering. It is one of the most fascinating and useful quantum features, being an important resource for quantum key distribution, randomness certification, subchannel discrimination, among others (see Ref. [3]). However, all these require, implicitly, the ability to discern which quantum states can be steered and which cannot.

This has proven to be a difficult task, even for the simplest scenario of two qubits. In their Msc. thesis [4], Gabriela Ruiz tried to estimate the relative size of the set of steerable two-qubit states. Among other things they trained a neural network approach to certifying steerability and were able to reach an accuracy of 89.4%.

In this work, we aim to improve upon these results in two main areas. First, we will enhance the training set by incorporating the state-of-the-art steering certification algorithm proposed by Ref. [5]. Second, we will explore the suggestion from Ref. [6] that leveraging convolutional networks can be beneficial when working with quantum states, which are inherently matrices and thus suitable for convolutional methods.

The classical approach typically involves using a dense real network that takes as input the coefficients in the Pauli basis. However, it may be more suitable to use a complex convolutional network. This approach aligns more naturally with the underlying mathematical structure of quantum states, since quantum states are already represented as complex matrices. To build these complex network, we will utilize the method proposed in Ref. [7], which also allows the construction of other hypercomplex networks, which could potentially offer additional benefits.

To evaluate the efficiency of these different networks, we developed three networks: a dense real network similar to the one implemented in Ref. [4], a convolutional complex network, and a convolutional quaternionic network. We trained all of them to certify steerability in two-qubit states. Our findings indicate that all networks performed similarly, achieving a balanced accuracy of 94%. However, the complex network achieved this performance with a slightly smaller number of parameters.

2. BACKGROUND

2.1. The basics

In quantum theory, the state of a system is described by a positive semidefinite operator ρ over a Hilbert space $\mathcal{H}$ satisfying $\mathrm{tr}(\rho) = 1$. The set of all such states is denoted as $\mathcal{S}(\mathcal{H})$.

For two systems A and B, associated with the Hilbert spaces $\mathcal{H}_A$ and $\mathcal{H}_B$, the composite system is associated with the space $\mathcal{H}_{AB} = \mathcal{H}_A \otimes \mathcal{H}_B$.

A composite state $\rho^{AB} \in \mathcal{S}(\mathcal{H}_A \otimes \mathcal{H}_B)$ is separable if it can be written as

$$\rho^{AB} = \sum_i \rho_i^A \otimes \rho_i^B,$$

for some $\rho_i^A \in \mathcal{S}(\mathcal{H}^A)$ and $\rho_i^B \in \mathcal{S}(\mathcal{H}^B)$. A state that is not separable is entangled.

Interestingly, when a state is entangled, we cannot attribute a state to each individual subsystem. In quantum theory, the whole can be more than the sum of its parts. The best description we can give to a subsystem is the reduced state. The reduced state of $\rho^{AB} \in \mathcal{S}(\mathcal{H}_A \otimes \mathcal{H}_B)$ with respect to a subsystem $\mathcal{H}_A$ is

$$\rho^A := \mathrm{tr}_B(\rho^{AB}),$$

where tr_B is the partial trace over the $\mathcal{H}_B$ subsystem. Note that, in general, $\rho^{AB} \neq \rho^A \otimes \rho^B$.

In quantum theory, a measurement x corresponds to an orthogonal decomposition of the Hilbert space $\mathcal{H} = \bigoplus \mathcal{H}_i$. Each subspace $\mathcal{H}_i$ is associated with an outcome a such that the probability of obtaining outcome a after measuring x on a system prepared at state ρ is

$$P_\rho(a|x) = \mathrm{Tr}(\Pi_x^a \rho),$$

where Π_x^a is the orthogonal projector on the $\mathcal{H}_i$ subspace. The set of all possible measurements acting on the state space $\mathcal{S}(\mathcal{H})$ is denoted by $\mathcal{M}(\mathcal{H})$.

After implementing measurement $x \in \mathcal{M}(\mathcal{H})$ on a state $\rho \in \mathcal{S}(\mathcal{H})$ and obtaining outcome a, the state of the system is updated to

$$\rho_{a|x} = \frac{\Pi_x^a \rho \Pi_x^a}{\mathrm{tr}(\Pi_x^a \rho)}.$$

In quantum computation, the most basic subsystem is the quantum bit (qubit). A system with n qubits is associated with the Hilbert space $\mathcal{H}_n = (\mathbb{C}^2)^{\otimes n}$, meaning that $\mathcal{S}(\mathcal{H}_n)$ corresponds to the space of $2^n \times 2^n$ complex matrices that are self-adjoint, positive semidefined, and have trace one (which is a real vector space). Measurements in this context can be associated, via the spectral theorem, to the set of $2^n \times 2^n$ self-adjoint complex matrices.

2.2. Quantum steering

As mentioned in the introduction, steering pertains to the ability of one party to steer the state of the other depending on the chosen measurement. To understand steering, let us consider two very distant observers, Alice and Bob, each controlling a different part of a system, which we will associate the the Hilbert spaces $\mathcal{H}^A$ and $\mathcal{H}^B$. Each measurement that Alice can implement in her system will be of the type $\mathcal{S}(\mathcal{H}^A) \otimes \mathbb{1}^B$, and similarly for Bob, since neither is able to directly interact with the far away system.

Suppose that their system is prepared on state ρ, and Alice implements a measurement $x \in \mathcal{S}(\mathcal{H}^A)$; Bob's reduced state will then be $\rho_{a|x}$ with probability $P_\rho(a|x)$ for each outcome a. We summarize this information in an assemblage

$$\sigma_{a|x} := \mathrm{tr}_A((\Pi_{a|x} \otimes \mathbb{1}\rho^{AB})),$$

which is a set of subnormalized reduced states.

After communicating with Alice and characterizing the assemblage $\{\sigma_{a|x}\}_{a,x}$, Bob may try to explain his statistics assuming his state was not influenced by Alice's measurements: He conjectures that initially his particle was in some latent state ρ_λ^B with probability $P(\lambda)$, parameterized by some latent variable λ. Then, Alice measuring and broadcasting her result result just gave him additional information on the hidden probability distribution. This description, implies the states have the form

$$\sigma_{a|x} = \sum_\lambda P(a|x,\lambda) P(\lambda) \rho_\lambda^B,$$

which we call a local hidden state model.

If for a given state $\rho \in \mathcal{S}(\mathcal{H}^A \otimes \mathcal{H}^B)$, every local measurement $x \in \mathcal{S}(\mathcal{H}^A) \otimes \mathbb{1}^B$ that Alice implements results in an assemblage that can be written in the form of Eq. (2.2), we say that ρ is an unsteerable state, otherwise we say it is steerable.

Certifying if a state is steerable is a difficult problem. All separable states are clearly unsteerable, but some entanglement states are also unsteerable, making the problem complex.

3. METHODS

3.1. Generating the states

To train our networks with an unbiased set, we used a database of states uniformly generated using the Bures metric. To define it, we first need to define the fidelity of two states

$$F(\rho_1, \rho_2) := \left(\mathrm{tr}\sqrt{\sqrt{\rho_1}\rho_2\sqrt{\rho_1}}\right)^2,$$

which can be interpreted as a sort of statistical distinguishability between states.

We can then define the Bures metric in the state space

$$D_B\left(\rho_1,\rho_2\right) := \sqrt{2\left(1-\sqrt{F\left(\rho_1,\rho_2\right)}\right)}.$$

To generate a uniformly random two-qubit state using the Bures metric, we take a 4×4 complex random matrix G from the Ginibre ensemble, a random unitary matrix U distributed according to the Haar measure on $U(4)$ and write

$$\rho = \frac{(\mathbb{1}+U)GG^{\dagger}\left(\mathbb{1}+U^{\dagger}\right)}{\operatorname{tr}\left[(\mathbb{1}+U)GG^{\dagger}\left(\mathbb{1}+U^{\dagger}\right)\right]}.$$

For this work, we utilized a large database of two-qubit states generated in this manner that is available on Github[1].

3.2. Certifying entanglement

In the case of a two-qubit system, it is very simple to determine whether a state is entangled or not. The the positive partial trace criterion (PPT) [8]. A state $\rho \in \mathcal{S}(\mathcal{H}^A\otimes\mathcal{H}^B)$ is entangled if an only if its partial transpose with respect to either subsystem

$$\rho^{T_B} = (I^A\otimes T^B)(\rho)$$

is positive defined. This is very simple to determine computationally, and can be done by calculating the SVD decomposition of the state.

3.3. Certifying steerability

The two-qubit system is the only system where the set of steerable states is completely characterized. The state-of-the-art algorithm for determining whether a state is steerable was proposed by Ref. [5] and it leverages geometric properties of state and measurement sets.

To certify that a state is steerable, one needs to optimize the Critical Radius function:

$$R(\rho^{AB}) = \min_{C,\mu}\frac{1}{\sqrt{2}\|\operatorname{Tr}_B[\bar{\rho}(\mathbb{1}^A\otimes C)]\|}\int_{\sigma\in\mathcal{B}_B} d\mu(\sigma)|\operatorname{tr}_B(C\sigma)| \tag{1}$$

where $\|\cdot\|$ is the trace norm, $\bar{\rho} = \rho^{AB} - (\mathbb{1}^A\otimes\rho^B)/2$, C is optimized over all measurements on $\mathcal{H}_B$ and μ over all distribution over the set of qubit state-space.

The only downside is that, to calculate the optimisation, we need to approximate the qubit state-space with a polytope, since this approximation is not perfect, there are a few states for which the method is inconclusive. In our analysis we utilized the polytope generated by solving the covering problem with 72 points on a sphere, which is available in Ref. [9].

[1] `https://github.com/BiduRuiz/SDP_LocalModels`

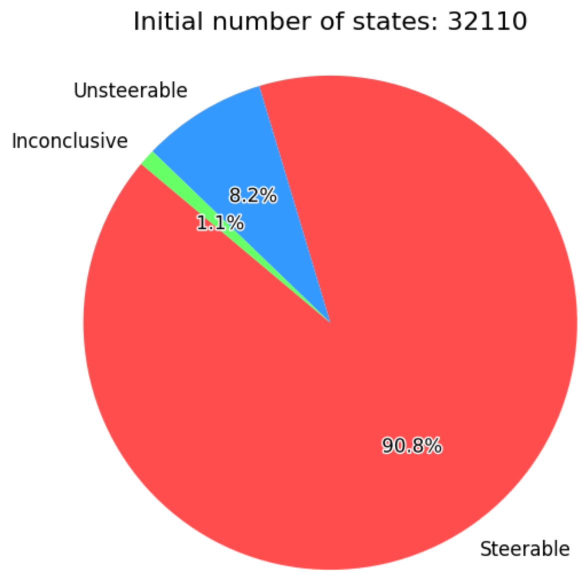

Fig. 1. Category breakdown of the generated states.

3.4. Neural networks

To implement our neural networks, we utilized the TensorFlow library in Python. Although TensorFlow supports complex numbers, it does not natively handle quaternions or other hypercomplex numbers. Therefore, for our complex and quaternionic networks, we utilized custom layers based on the work of Ref. [7]. In this reference, the authors explain how to implement such layers by separating the real and imaginary parts and using only real operations, which allows us to leverage the benefits of automatic differentiation provided by TensorFlow. The names of the layers we used are ComplexConv, ComplexDense, QuatConv, and QuatDense, which are respectively for complex and quaternionic numbers. All the algorithims utilized in this project are accesible in our Github repository[2].

4. RESULTS

After generating the states using the Bures metric and certifying entanglement with the PPT criterion, we were left with 32110 entangled two-qubit states. We them characterised them with respect to steerability, following the strategy from Sec. 3.2, obtaining the breakdown displayed in Fig. 1.

Clearly we have a very unbalanced set, which was something that needs to addressed before training. After removinf the inconclusive states, we separated 10% of the valid states

[2] `https://github.com/turdutra/Steerability-with-NN`

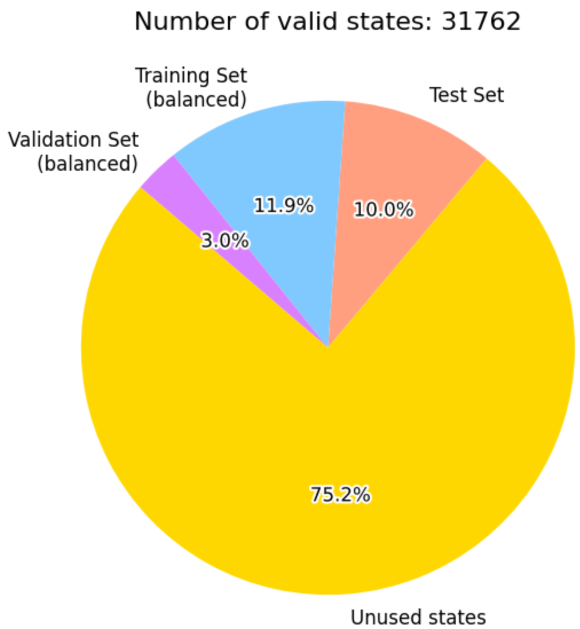

Fig. 2. Visualization of our training, test, and validation sets

	Real	Complex	Quaternionic
No. of real parameters	122219	82166	114910

Table 1. Number of real parameters for each model

as our test set. We balanced the remaing states and took 80% os the balanced set as out training set and 20% as out validation set. The full distribution can be visualized in Fig.2.

To prepare the inputs for the networks, we employed different methods for each type. For the real model, we calculated the 15 coordinates of the matrices in the Pauli basis. For the complex model, we used the state matrix as it is. For the quaternionic model, we added two layers of zeros to the state matrix to create a quaternionic matrix.

When designing the networks, we chose to establish that all of them would have the following architecture: two intermediate layers with a dropout layer between them. To tune the hyperparameters of those layers, we used the Keras Tuner library. We configured it to tune the number of neurons, filters, and dropout rate, and set a maximum size of around 240,000 real parameters for all networks.

Fig. 3 illustrates the architecture and displays the number of parameters of each layer in network after tuning. Table 1 computes the total number of parameters for each model. The complex network stands out for being noticeably smaller than the others.

We trained all three networks for 100 epochs, using a batch size of 32 and implementing a step-decay learning rate. Fig 4 shows the accuracy in the validation set for each network during training as a function of the epoch. All networks achieved very high accuracy values, with the quaternionic network outperforming the others in this metric.

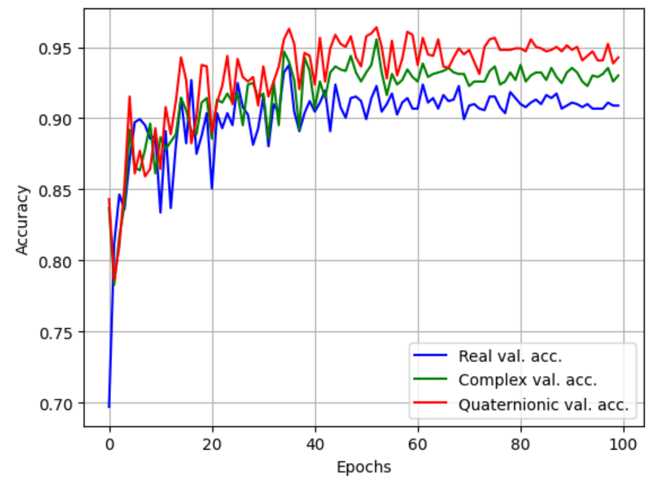

Fig. 4. Validation accuracy for each network during training

Of course the true measure of successes for machine learning networks is the accuracy in the test set, which is completely uninvolved in the training process. Fig 5 presents the confusion matrix and the balanced accuracy of all three networks in the test set. We find that all networks exhibit remarkably similar balanced accuracy.Notably, the balanced accuracy of the real-valued network on the test set surpasses the accuracy on the validation set. This suggest that the validation set may be too small, and not fully representative of the entire dataset.

5. CONCLUSION

In this study, we successfully employed convolutional neural networks to certify the steerability of quantum states. Our approach using a diverse set of two-qubit states generated uniformly from the Bures metric and training various neural network architectures, including real dense, complex convolutional, and quaternionic convolutional networks to detect steerability of a two-qubit state. All networks achieved high accuracy, with balanced accuracies of 94% on the test set. Notably, the complex convolutional network achieved comparable performance with fewer parameters, highlighting its efficiency.

It is important to note that while the network method may not be as reliable as the method proposed by Ref.[5] for certifying the steerability of a single state, it can be highly useful for creating large datasets of either steerable or entangled unsteerable states. The network can quickly preselect states of a

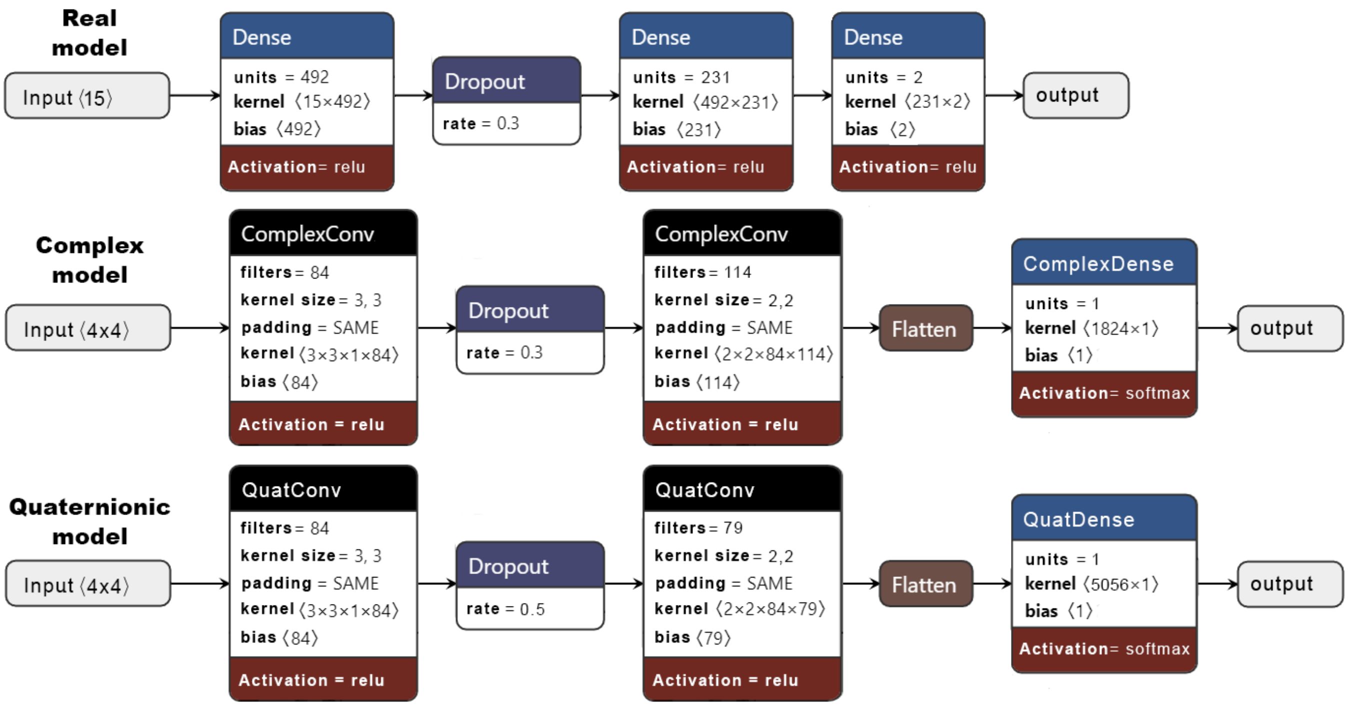

Fig. 3. Structure and hyperparameters utilized for each model. Note that the units in the complex network correspond to complex numbers, and, analogously, for the quaternionic model.

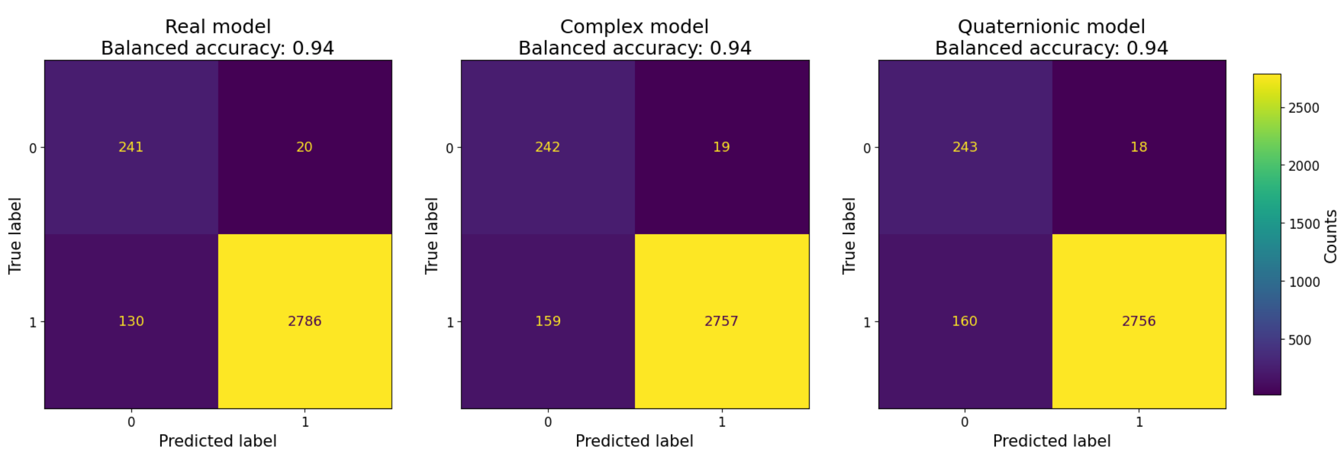

Fig. 5. Balanced accuracy for each network in the test set

specific type, allowing the more precise method by Ref.[5] to be applied to a smaller subset for confirmation, thereby optimizing the process.

Looking forward, the success of convolutional networks in this context opens up exciting possibilities. Similar to their application in image processing, which does not require a static picture size, convolutional networks could be trained to identify steering and other properties in quantum states of different dimensions. A network capable of processing quantum states of varying dimensions would be a powerful tool, and currently, there is available tool with such capabilities. This approach could significantly enhance our ability to analyze and certify complex quantum systems.

Acknowledgements

We would like to thank Lucas Porto and Gabriela Ruiz, for the insights and the assistance in building the dataset. We would also like to thank Prof. Marcos Eduardo Valle for the lectures on how to build the custom layers and the support for writing this paper. This study was financed in part by the Coordenação de Aperfeiçoamento de Pessoal de Nível Superior (CAPES)- Finance Code 001.

6. REFERENCES

[1] Albert Einstein, Boris Podolsky, and Nathan Rosen, "Can quantum-mechanical description of physical reality be considered complete?," *Physical review*, vol. 47, no. 10, pp. 777, 1935.

[2] E. Schrödinger, "Discussion of probability relations between separated systems," *Mathematical Proceedings of the Cambridge Philosophical Society*, vol. 31, no. 4, pp. 555–563, 1935.

[3] Roope Uola, Ana C. S. Costa, H. Chau Nguyen, and Otfried Gühne, "Quantum steering," *Rev. Mod. Phys.*, vol. 92, pp. 015001, Mar 2020.

[4] Gabriela Ruiz, "Local entangled states certification," Master's dissertation, Universidade Estadual de Campinas (UNICAMP). Instituto de Física Gleb Wataghin, Campinas, SP, August 4 2023, Examining board members: Rafael Luiz da Silva Rabelo, Marcelo de Oliveira Terra Cunha, Leonardo Guerini de Souza.

[5] H. Chau Nguyen, Huy-Viet Nguyen, and Otfried Gühne, "Geometry of einstein-podolsky-rosen correlations," *Phys. Rev. Lett.*, vol. 122, pp. 240401, Jun 2019.

[6] Ali Kookani, Yousef Mafi, Payman Kazemikhah, Hossein Aghababa, Kazim Fouladi, and Masoud Barati, "Xpookynet: Advancement in quantum system analysis through convolutional neural networks for detection of entanglement," 2024.

[7] Chiheb Trabelsi, Olexa Bilaniuk, Ying Zhang, Dmitriy Serdyuk, Sandeep Subramanian, João Felipe Santos, Soroush Mehri, Negar Rostamzadeh, Yoshua Bengio, and Christopher J Pal, "Deep complex networks," 2018.

[8] Asher Peres, "Separability criterion for density matrices," *Phys. Rev. Lett.*, vol. 77, pp. 1413–1415, Aug 1996.

[9] R. H. Hardin, N. J. A. Sloane, and W. D. Smith, "Tables of spherical codes with icosahedral symmetry," `http://neilsloane.com/icosahedral.codes/index.html`, 2012, Accessed: 2024-06-13.