Super-additivity of noisy quantum channels with the help of Neural network ansatz and Genetic Algorithm

A Project Report Submitted

in Partial Fulfilment of the Requirements

for the

MINOR DEGREE

in

DATA SCIENCE

by

ANANTHA KRISHNAN S
ROLL NO IMS18022

to

SCHOOL OF MATHEMATICS
INDIAN INSTITUTE OF SCIENCE EDUCATION AND RESEARCH
THIRUVANANTHAPURAM - 695 551, INDIA

July 2022

DECLARATION

I, **Anantha Krishnan S (Roll No: IMS18022)**, hereby declare that, this report entitled **"Super-additivity of noisy quantum channels with the help of Neural network ansatz and Genetic Algorithm"** submitted to Indian Institute of Science Education and Research Thiruvananthapuram towards partial requirement of **Minor Degree** in **Data Science**, is an original work carried out by me under the supervision of Dr. Nagaiah Chamakuri 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**

July 2022

CERTIFICATE

This is to certify that the work contained in this project report entitled **"Super-additivity of noisy quantum channels with the help of Neural-network ansatz and Genetic Algorithm"** 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 | Dr. Nagaiah Chamakuri

July 2022 | Project Supervisor

ACKNOWLEDGEMENT

I want to extend a sincere and heartfelt obligation towards all the personages without whom the completion of the project was not possible. I express my profound gratitude and deep regard to Dr. Nagaiah Chamakuri, IISER Thiruvananthapuram for his guidance, valuable feedback, and constant encouragement throughout the project. His valuable suggestions were of immense help. I sincerely acknowledge his constant support and guidance during the project.

I am immensely grateful to Govind Lal Siddharth and Dr. Manik Banik for their constant support and encouragement. I am also grateful to the Indian Institute of Science Education and Research, Thiruvananthapuram, for allowing me to do this project and providing all the required facilities.

Thiruvananthapuram - 695 551 **Anantha Krishnan S**

July 2022

ABSTRACT

Machine learning has greatly improved the attainability of solutions when it comes to challenging problems such as those in fundamental research. Here, we use machine learning to tackle a quantum information scenario. Quantum channels exhibit a convenient property called superadditivity which can improve transmission rates in quantum information channels. In this work we found quantum states expressing Superadditivity using neural network variational ansatz and optimization algorithms; through this we explored the usefulness of using neural network as a variational state ansatz for representing Quantum qubit states in the context of quantum information-processing tasks, and also the efficiency of using different optimization algorithms to find the better case. (a) Neural network states yield quantum codes with a high Coherent-information for Qubit pauli channels, Dephrasure channel and De polarizing channel; such codes have proven to outperform all other known codes for these channels. (b) It has also been shown that the quantum codes for single channel cases were given by repetition codes. (c) Using genetic algorithm has proved to be more effective than using simple gradient-based methods for optimizing the neural network ansatz representing the states.

Contents

Chapter 1

Introduction

When it comes to the usefulness of using a particular channel for communication, we use its capacity as a tool for understanding this parameter. In the case of classical channels, we find the capacity, which is given by the mutual information between input and output random variables. This mutual information is an entropic quantity that is optimized over all possible input probability distributions and hence is the property of the channel. In classical mechanics, there also exists the intuitive additive property, which makes parallel uses of many identical copies of a channel equivalent to multiple uses of its single-copy.

While modeling the quantum communication scenario, classical input and output variables are replaced by respective input and output Hilbert spaces, and accordingly, the channel action is most generally described by a completely positive trace preserving (CPTP) map from the operator space on input to operator space on output Hilbert space.

The capacity of a channel quantifies how good it is to transfer information, classical and quantum information when it comes to classical and quantum channels respectively. The entropic quantity - coherent information maximized over all quantum states, defines the quantum capacity.

Unlike the additive property of a classical channel, the quantum capacity of a quantum channel can exhibit a striking super-additive phenomenon, where two different channels each with zero quantum capacity can be used together to send quantum information with nonzero rate. Later, it has been shown that there exists a quantum channel whose N-parallel use is no good to send quantum information whereas its (N + 1)-parallel use becomes useful for transferring quantum information, with N being arbitrary integer values. This super-additivity feature can make the parallel use of many copies of channel more beneficial than multiple uses of a single channel.

The phenomenon of super-additivity has been demonstrated in three-parameter family of qubit pauli channels for both two-channel use and three-channel use, in the one-parameter family of Depolarising channels both three-channel use and five-channel use displayed superadditivity. More recently, super-additivity of coherent information was demonstrated for two channel-uses for the two-parameter family of Dephrasure Channels again using the repetition code.

In the project, we utilize machine learning-based optimization techniques to check each of its efficiency when it comes to finding quantum codes expressing superadditivity of coherent information. We explored the usefulness of using neural networks as a variational ansatz for representing quantum qubit states and study the problems encountered while using gradient-descent as an optimization technique, and the efficiency of genetic algorithms for the same purpose. Specifically, we use genetic algorithms on the neural network ansatz to test out regions of three-parameter space of qubit pauli channel for superadditivity of quantum codes.

Local Hamiltonians with highly-entangled ground states only require a polynomial number of parameters to describe, as do quantum circuits of polynomial depth. This fact motivates the use of variational representations of quantum states to solve a large class of problems. At the heart of any variational ansatz is the idea to preserve as much information about the quantum state as possible, while discarding irrelevant features.

1.1 Main results

(a)Neural networks work efficiently as a variational ansatz for representing quantum qubit states (b)For single channels maximum capacity for quantum communication was shown by repition codes (c)Found genetic algorithms to be pretty effective in optimizing the neural network ansatz to the state having maximum coherent information for qubit pauli channels (d)Simple gradient descent algorithms seem to get stuck at local minima on using negative coherent information as the loss function. (e)Confirmed superadditive nature in pauli qubit channels using genetic algorithms in qubit pauli channels

1.2 Structure of the paper

In Chapter 2, the work discusses the preliminaries required to understand the project. Section 2.1 and 2.2 discusses the state and the quantum mechanical problem of exponential growth. In Section 2.3 capacity for classical communication and quantum communication and the mathematical representation of coherent information is explained in detail. It is also explained how quantum capacity is an achievable rate. The property of superadditivity and qubit channels are also explained in detail. Chapter 3 introduces the machine learning techniques that are being used. The work uses a neural network as the variational ansatz similar to how it is in the Restricted Boltzmann Machine. Here both gradient descent and the genetic algorithm method were used to handle the optimization problem. The machine learning method that is being used is also explained in detail. In Chapter 4, the results and future directions have been given.

Chapter 2

Preliminaries

In this section, we will first review the idea of super-additivity of coherent information of a quantum channel, and then briefly discuss some concepts in machine learning that will be relevant to our purpose.

2.1 Density Operator

The quantum state of a physical system is described by a density matrix. It is a generalization of the more usual state vectors or wavefunctions: while those can only represent pure states, density matrices can also represent mixed-state. It allows for the calculation of the probabilities of the outcomes of any measurement performed upon this system, using the Born rule.

$$\rho = \sum_j p_j |\psi_j\rangle\langle\psi_j| \quad , \quad \langle\psi_m|\psi_n\rangle = \delta_{mn} \quad , \quad \sum_n p_n = 1$$

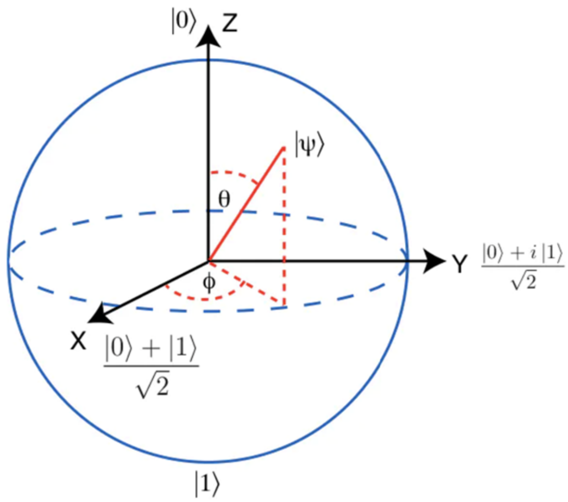
where ψ_j's are pure states prepared with probability p_j, known as an ensemble.

We are dealing with pure states, so we can represent density operators as $\rho = |\psi_{AB}\rangle\langle\psi_{AB}|$

2.2 Quantum Information and the problem of exponential growth

Quantum information is the information about the state of a quantum system. It is the basic entity of study in quantum information theory and can be manipulated using quantum information processing techniques. It is an interdisciplinary field involving quantum mechanics, computer science, information theory, philosophy, and cryptography. Its study is also relevant to cognitive science and neuroscience disciplines. Its main focus is on extracting information from matter at the microscopic scale.

2.2.1 Qubit and Bloch Sphere



Bloch Sphere

Quantum information differs strongly from classical information, epitomized by the bit, in many striking and unfamiliar ways. While the fundamental unit of classical information is the bit, the most basic unit of quantum information is the qubit. Classical information is measured using Shannon entropy, while the quantum mechanical analogue is Von Neumann entropy. Given a statistical ensemble of quantum mechanical systems with the density matrix ρ, it is given by $S(\rho) = -Tr(\rho ln \rho)$

The state of a qubit contains all of its information. This state is frequently expressed as a vector on the Bloch sphere. This state can be changed by applying linear transformations or quantum gates to them.

A qubit is a two-state or two-level quantum-mechanical system, one of the simplest quantum systems displaying the peculiarity of quantum mechanics. In a classical system, a bit would have to be in one state or the other. However, quantum mechanics allows the qubit to be in a coherent superposition of both states simultaneously, a property that is fundamental to quantum mechanics and quantum computing. Unlike classical digital states which are discrete, a qubit is continuous-valued, describable by a direction on the Bloch sphere.

2.2.2 Problem of exponential growth

A system consisting of single qubits has a computational basis of 2 elements; $|0\rangle$ and $|1\rangle$. A two-qubit system likewise will have 4 computational basis elements; $|00\rangle, |01\rangle, |10\rangle$ and $|11\rangle$. So a N qubit system will have a computational basis of 2^N elements; $|00\ldots00\rangle, |00\ldots01\rangle ... |11\ldots11\rangle$.

Associated with each computational basis state there is a probability amplitude α_i which is a complex number that determines what kind of quantum state it is.

$$|\phi> = \alpha_0|000\rangle + \alpha_1|001\rangle + \ldots + \alpha_7|111\rangle$$

for three qubits.

Here as the number of particles increases the dimension of the problem increases exponentially as 2^N , $N =$ no of particles. For every N qubits we have to study, the density matrix will have a power of $2^N \times 2^N$; add into this a purified state, we will be looking at $2^{2N} \times 2^{2N}$ order density matrices. So its necessary to use approaches with meta-heuristic algorithms.

2.3 Quantum Information Theory

Information theory is based on probability theory and statistics. Information theory often concerns itself with measures of information of the distributions associated with random variables.

Important quantities of information are entropy, a measure of information in a single random variable, and mutual information, a measure of information in common between two random variables. Entropy - The former quantity is a property of the probability distribution of a random variable and gives a limit on the rate at which data generated by independent samples with the given distribution can be reliably compressed. Mutual Information - The latter is a property of the joint distribution of two random variables and is the maximum rate of reliable communication across a noisy channel within the limit of long block lengths when the channel statistics are determined by the joint distribution.

The classical theory of information is explained using the studies pioneered by Shannon. A noisy classical channel is a stochastic map $\mathrm{N} : \mathrm{P(X)} \rightarrow \mathrm{P(Y)}$, where P(X) and P(Y) respectively denote probability distributions on the input and output random variables X Y.

A key goal of quantum information theory is to extend the classical theory of information, as pioneered by Shannon, to include quantum effects like superposition and entanglement. The capacity of a noisy communication channel plays a fundamental role in classical information theory: it is the optimal noiseless communication rate that a noisy channel can support. In the quantum setting, a noisy communication channel has multiple capacities since it can be used to accomplish different communication tasks. Thus, a quantum channel N has a capacity for classical communication C(N), quantum communication Q(N), and private classical communication P(N)

2.3.1 Capacity for classical communication

In quantum information theory, the classical capacity of a quantum channel is the maximum rate at which classical data can be sent over it error-free in the limit of many uses of the channel. The capacity of a classical channel $N : X \to Y$ is given by

$C(N) = C^{(1)}(N) = max_X I(X : Y)$, where the maximization is over all input probability distributions and the mutual information. (It quantifies the amount of information obtained about one random variable by observing the other random variable) $I(X : Y) = H(X) + H(Y) - H(XY)$

Shannon entropy $H(X) := \sum_{x \in X} p(x) log(p(x))$

There exist an intuitive property for the capacity of classical communication called additivity. Using N classical channels parallelly to send information is the same as sending information through the classical channels N times.

$C^{(1)}(N^{\otimes}) = nC^{(1)}(N)$

$C(N) = lim_{n->\infty}(1/n)C^{(1)}(N^{\otimes})$ This is the regularized formula. But, when additivity is introduced regularization disappears $C(N) = C^{(1)}(N)$

The additive property greatly simplifies the optimization problem over an unbounded number of channel uses.

2.3.2 Capacity for quantum communication

In this case the classical input and output variables are replaced by respective input and output Hilbert spaces, and accordingly, the channel action is most generally described by a completely positive trace preserving (CPTP) map from the operator space on input to operator space on output Hilbert space.

The channel action on a Quantum state is a quantum operation and it has to be a Completely Positive Trace Preserving map.
Complete Positivity: $(\Lambda_A \otimes I_R)\rho_{AR} \geq 0 \;\; \forall \;\; \rho_{AR} \;\; \in \;\; P(H_A \otimes H_R)$
Trace preserving: $Tr(\Lambda_z(\rho)) = 1$

Quantum channels do quantum operations on the input quantum state In the case of completely positive maps, we can model quantum operations for quantum states using the Kraus operators/noise operators.

$$\phi(\rho) = \sum_k B_k \rho B_k^*$$
$$\sum_k B_k^* B_k = 1$$

Coherent information is the entropy measure used in Quantum information theory. It attempts to describe how much quantum information in the state will remain after it passes through a quantum channel. In this sense, it is intuitively similar to the mutual information of classical information theory. The coherent information is written $I(\rho, N)$

$I(\rho, N) = S(N\rho) - S(N, \rho)$ where $S(N\rho)$ is the von Neumann entropy of the output and $S(N, \rho)$ is the entropy exchange between the state and the channel.

$$S(\phi, \rho) = S[Q', R'] = S(\rho'_{QR})$$

where $S(\rho'_{QR})$ is the von Neumann entropy of the system Q and a fictitious purifying auxiliary system R after they are operated on by ϕ.

The quantum capacity of a quantum channel captures its utility to transfer quantum information. The quantum capacity $Q(\Lambda)$ of a quantum channel Λ is given by the following regularized expression.

$$Q(\Lambda) = \lim_{n\to\infty} \frac{1}{n} Q^{(1)}(\Lambda^{\otimes n})$$

$Q^{(1)}(\Lambda^{\otimes n})$ is the n-shot channel capacity or (channel coherent information), where n identical copies of Λ are used in parallel. Q Capacity formula involves the evaluation of the channel coherent information over an unbounded no of channels.

$$Q^{(1)}(\Lambda^{\otimes n}) = max_{\rho_S} I(\rho_S,^{\otimes n})$$

Channel coherent information or n-shot channel capacity is Coherent information maximized for all quantum states.

$$\begin{aligned} \rho_S &\in D(H^{(\otimes n)}) \\ \rho_S &\in Tr_R(|\psi_{SR}\rangle\langle\psi_{SR}|) \\ I(\rho_S, \Lambda^{\otimes n}) &= I(|\psi_{SR}\rangle, \Lambda^{\otimes n}) \end{aligned}$$

Expression for Coherent information of $\Lambda^{\otimes n}$ channel operation for the input ρ_S is:

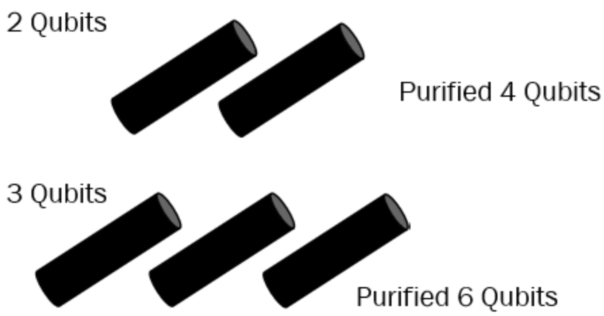
$$I(|\psi_{SR}\rangle, \Lambda^{\otimes n}) = S(\Lambda^{\otimes n}(\rho_S)) - S(\Lambda^{\otimes n} \otimes I_R(|\psi_{SR}\rangle\langle\psi_{SR}|))$$

S is Von-Neumann entropy

$S(\rho) = -Tr(\rho log(\rho))$

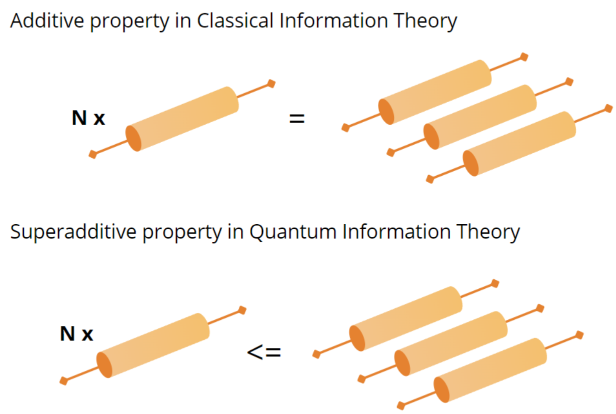
Where ρ is the density matrix

2.3.3 Purification



In a state ρ_A of a quantum system A, it is possible to introduce another system, which we denote R, and define a pure state $|AR\rangle$ for the joint system AR such that $\rho_A = Tr_R(|AR\rangle\langle AR|)$ That is, the pure state $|AR\rangle$ reduces to ρ_A when we look at system A alone. This is a purely mathematical procedure, known as purification, which allows us to associate pure states with mixed states. For this reason, we call system R a reference system: it is a fictitious system, without a direct physical significance. If we want to send n qubits through a 2-shot quantum channel, another n qubits are in the reference /environment system. So in total For the $|\psi_{SR}>$ state there will be 2n qubits

2.3.4 Superadditivity



Earlier additivity simplified the Classical communication scenario, but that is not the case in the Quantum scenario. When few channels show weak additivity

$$Q^{(1)}(N^{\otimes n}) \ \leq nQ^{(1)}(N)$$
$$Q(\Lambda) = \lim_{n\to\infty} \tfrac{1}{n} Q^{(1)}(\Lambda^{\otimes n}) \to Q(\Lambda) = Q^{(1)}(\Lambda)$$

the regularisation disappears.

But there are channels with strictly super-additive channel coherent information. Where the Quantum capacity when using N-Quantum channels parallelly (simultaneously) is greater than when used N time individually.

$$Q^{(1)}(\Lambda^{\otimes n)}) > nQ^{(1)}(\Lambda)$$

For low noise channels the effect is low, so we take high noise regime. But here the regularisation is necessary. So

$$Q(\Lambda) = \ lim_{n\rightarrow\infty}\frac{1}{n}Q^{(1)}(N^{\otimes n})$$

Here we cant get rid of the unbounded no of channel copies

Despite the downside of having to do the optimization over an unbounded number of channel copies. We have another way channel coherent information is an achievable rate.

$$Q(\Lambda) \geq \frac{1}{n}Q^{(1)}(\Lambda^{\otimes n}) \geq \frac{1}{n}I(|\psi_{SR}\rangle, \Lambda^{\otimes n})$$

For a purified state $|\psi_{SR}\rangle$. So by maximizing the $\frac{1}{n}I(|\psi_{SR}>, \Lambda^{\otimes n})$ over all possible Quantum states we can achieve the Quantum capacity. And by comparing this achieved Channel capacity we can find states that express superadditivity.

The advantage of superadditivity is greater transmission rates in communication. Parallel use of multi-copy channels with properly constructed quantum codes can yield more transmission rates than using the channel multiple times.

2.3.5 Quantum Channels

The Quantum channel is a quantum operation that does an operation on a Quantum state.

Quantum Channel $\phi(\rho) = \sum_k B_k \rho B_k^*$, $\sum_k B_k^* B_k = I$

The B_k are Kraus operators and the action of the quantum channel can be represented like this.

- Qubit Pauli Channels

- Generalized Amplitude Damping Channel
- Dephrasure Channel
- Depolarizing Channel

Qubit Pauli Channel

$\Lambda(\rho) = p_0\sigma_0\rho\sigma_0 + p_1\sigma_1\rho\sigma_1 + p_2\sigma_2\rho\sigma_2 + p_3\sigma_3\rho\sigma_3$

$\sigma_0, \sigma_1, \sigma_2$ and σ_3 are pauli matrices. And $p_i \geq 0 \ \forall \ i$, $\sum\limits_{i=0}^{3} p_i = 1$ and $\rho \in D(\mathbb{C}^2)$

$A_p = \sum_{i=1}^{4} A_i \rho A_i^\dagger$

$A_1 = \sqrt{p_0}(|0\rangle\langle 0| + |1\rangle\langle 1|)$

$A_2 = \sqrt{p_1}(|0\rangle\langle 1| + |1\rangle\langle 0|)$

$A_3 = \sqrt{p_2}(-i|0\rangle\langle 1| + i|1\rangle\langle 0|)$

$A_4 = \sqrt{p_3}(|0\rangle\langle 0| - |1\rangle\langle 1|)$

Depolarizing Channel

$D_p(\rho) = (1-p)\rho + ptr(\rho)(|0\rangle\langle 0| + |1\rangle\langle 1|)/2$

Generalized Amplitude Damping Channel (GADC)

It is defined in terms of two parameters $p, \gamma \in [0, 1]$

$A_{\gamma,N} = \sum\limits_{i=1}^{4} A_i \rho A_i^\dagger$

$A_1 = \sqrt{1-N}(|0\rangle\langle 0| + \sqrt{1-\gamma}\ |1\rangle\langle 1|)$

$A_2 = \sqrt{\gamma(1-N)}|0\rangle\langle 1|$

$A_3 = \sqrt{N}(\sqrt{1-\gamma}|0\rangle\langle 0| + |1\rangle\langle 1|)$

$A_4 = \sqrt{\gamma N}|1\rangle\langle 0|$

Dephrasure Channel

It is defined in terms of two parameters $p, q \in [0, 1]$

$N_{p,q}(\rho) = (1-q)((1-p)\rho + p\sigma_3\rho\sigma_3) + qtr(\rho)|e\rangle\langle e|$

Chapter 3

Deep Learning and Genetic Algorithm

3.1 Deep Learning

At its core, machine learning techniques attempt to find patterns hidden in data and to exploit these learned patterns in control and decision making. With the advent of deep learning, complemented by the ever-increasing computational power, machine learning approaches have solved problems that were once thought to be impossible. Apart from this, deep learning techniques have shown exceptional performance in problems like computer vision, natural language processing, medical diagnosis, and more. Nevertheless, there is still a long way to go before the technology can achieve what is called Artificial General Intelligence. The success of modern machine learning in such diverse fields has prompted researchers to apply these techniques to fundamental research. Not surprisingly, many physicists have applied machine learning tools in their niche as well. Interestingly, there have also been attempts to develop methods for machine-assisted discovery of physical principles from experimental data.

Deep learning algorithms are machine learning models inspired by the structure and functioning of the brain. In most deep learning algorithms, the function to be learned is represented as a Deep Neural Network (DNN), an Artificial Neural Network (ANN) with many hidden layers. Although ANN architectures vary diversely across applications, certain fundamental features stay the same: The building units of ANNs are called neurons; these neurons are arranged into layers and connected together with weight parameters. A single neuron may receive multiple inputs, the weighted sum of which is passed through a non-linearity called the activation function. The output from the non-linearity is the output of the neuron, which is passed to other neurons or given as the output of the ANN

ANNs are universal function approximators; it was shown that with just one hidden layer and a non-linear activation function, an ANN with an appropriately chosen width could approximate any continuous function to arbitrary precision. Similar universal approximation theorems are also known DNNs but under slightly restricted conditions. Universal approximation capability appended with the back-propagation algorithm makes DNN an ideal model for learning highly complex patterns given a sufficient amount of data; moreover, DNNs have shown good generalization properties. In the past decade, the most exciting developments in AI involved deep learning, be it AlphaGo or self-driving cars. Unfortunately, DNNs are not foolproof; research has shown that poisoning training data sets can easily fool DNNs into making erroneous predictions. A lot of current AI research focuses on understanding this black box algorithm and making it foolproof.

3.2 Genetic Algorithm

Genetic algorithms (GA) are a class of meta-heuristic search/optimization algorithms inspired by the process of evolution by natural selection. All GAs start with a population of candidate solutions that are evolved over multiple generations to create better (fitter) solutions to the optimization problem. To simulate the various processes in natural selection, there are several biologically inspired operations involved in every GA. The typical GA begins with a randomly initialized set of candidate solutions. The candidate solutions are thought of as individuals of a population. Each individual is assigned a fitness that measures the quality of the individual. The population is left to evolve to produce fitter individuals and hopefully generate a good enough solution to the optimization problem. In any GA, the evolution process involves the following three fundamental steps.

- Cross-over/recombination: Several individuals of the population are selected and are crossed over by an appropriately chosen scheme to produce new individuals. Individuals with high fitness are more likely to be crossed over. The intuition is that crossing over can bring together useful features of fit individuals to produce fitter individuals.
- Mutation: The new individuals produced through cross-over are subjected to mutation. Mutation introduces more variations into the population.
- Selection: The fitter individuals are promoted to the next generation while individuals with less fitness are removed from the population with high probability.

GAs are derivative-free optimization methods, making these particularly useful for optimizing discontinuous objective functions and landscapes filled with local minima. Since GAs are good at avoiding local optima in complex optimization landscapes, GAs are often the tool of choice for global optimization problems. All this said, GAs have their limitations. Since information like the gradient is not considered in the optimization process, GAs can be resource hungry, requiring large population sizes and many fitness function evaluations. This can result in a substantial computational overhead if the fitness function evaluation is computationally expensive. In such cases, one often has to resort to parallel computing or computationally efficient approximations to the fitness function.

3.3 Principle

Representing Quantum state and Optimization
The Quantum state $|\psi_{SR}\rangle$ is represented in the Neural network ansatz depending on the weight matrices between the neural network's hidden layers. We initially take a large population of such sets of weight matrices of neural network states. We then use the optimization technique to find the value of coherent information of each individual in the population. In genetic algorithm, we will be taking the negative of coherent information as the loss function that is optimized to minima, and based on this we define the fitness of the individuals. After several generations, the optimization brings the weight matrices to a quantum state that gives the highest coherent information. Using this we will get the quantum channel and quantum code that gave the maximum coherent information. As explained earlier the maximum of coherent information gives the Quantum capacity. And comparing the data from single and multiple channel uses, we can find quantum channels that express superadditive property.

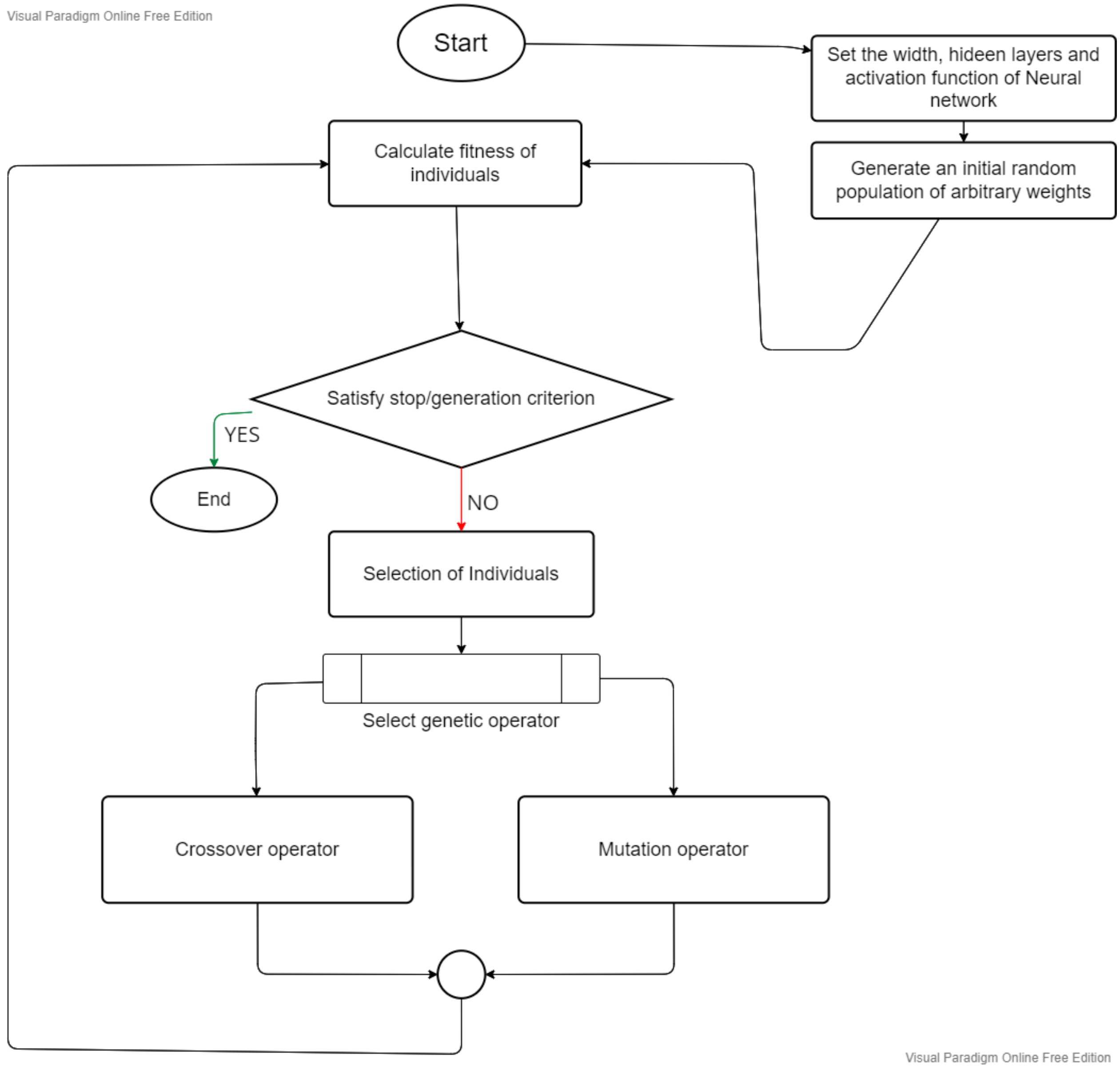


Figure 3.1: Neural network ansatz and Genetic Algorithm Flowchart

Machine Learning part

Machine(Neural network): We build a neural network with $w_1, w_2 \dots w_p$ as the weights and S computational basis as the input.
Non-linear function $\psi(S; w_1, \dots, w_p)$
The fully connected deep neural network architecture was implemented using the Sequential API of TensorFlow package

Learning(Optimization): The $w_1, \dots, w_p$ parameters are learned by minimizing the loss function i.e negative of coherent information, and satisfying the density operator conditions
$W = w_1, \dots, w_p$
Conditions of density operator: $tr(\rho) = 1$, $Tr(\rho^2) \leq 1$, $eigenvalues > 0$
We begin by setting a training vector, then the neural network ansatz represents a training quantum state. Then we find coherent information, optimize the loss function and take the output of the optimization technique and input it into the weights, thereby neural network represents a new quantum state. Through this we find the maximum coherent information and the corresponding Quantum states.

A pure quantum state can be expanded in the computational basis as

$$|\psi\rangle = \sum_{\{i^n\}} \frac{1}{C} \psi(\{i^n\}) |\{i^n\}\rangle$$

$\psi(\{i^n\})$ is the complex amplitude
$|\{i^n\}\rangle$ is the computational basis state identified by bit string $\{i^n\}$
C is the normalization factor

The simple genetic algorithm comes prebuilt with the DEAP package as deap.algorithms.eaSimple

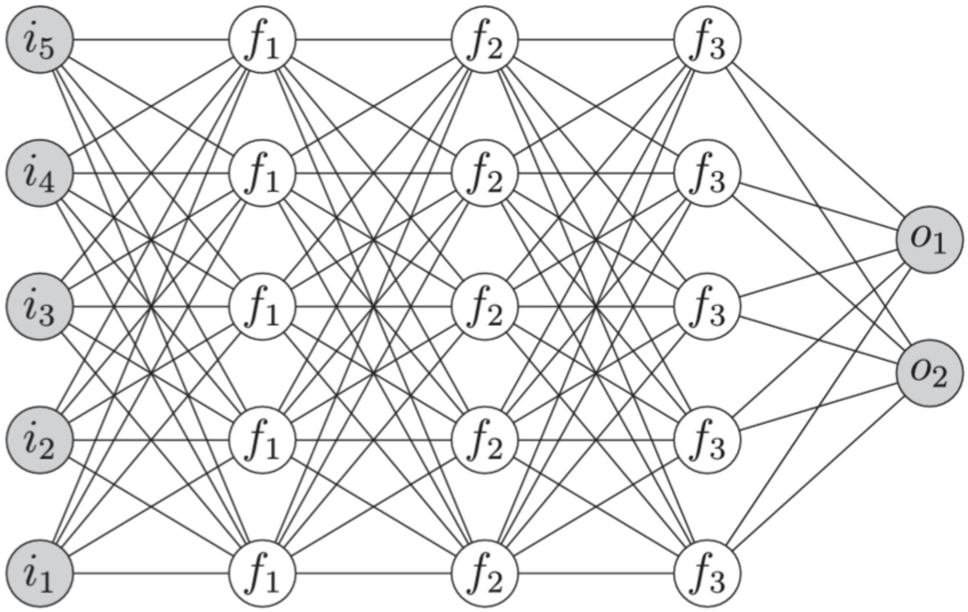


Figure 3.2: Feed forward Neural network

Advantages of using Neural network ansatz (Feedforward Neural network)

•Impressive accuracy

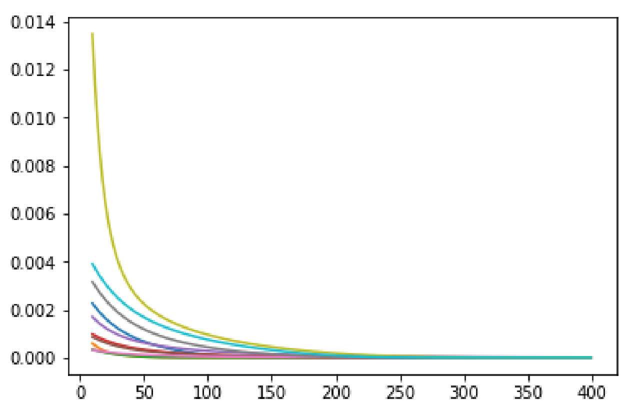
•Overcoming fidelity of quantum states

Chapter 4

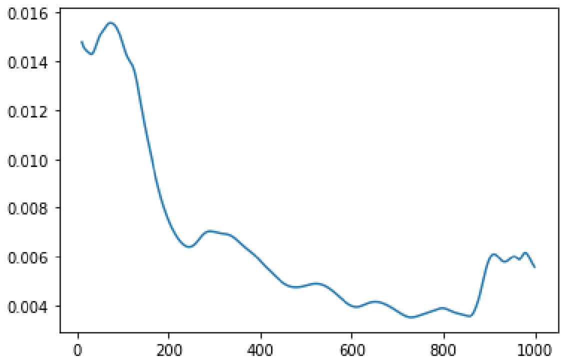
Results and Future Direction

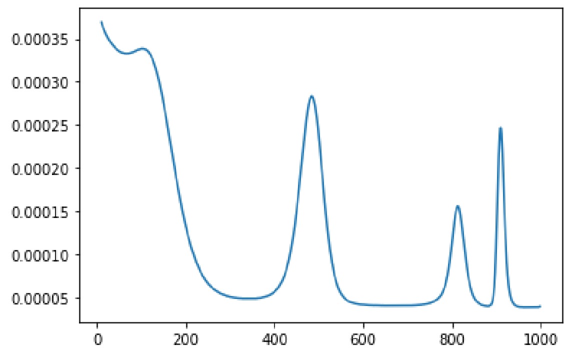
4.1 Results

- On using gradient descent from the tensorflow package as the optimization technique accurate results were obtained for single channel use in dephrasure channel and qubit pauli channels.

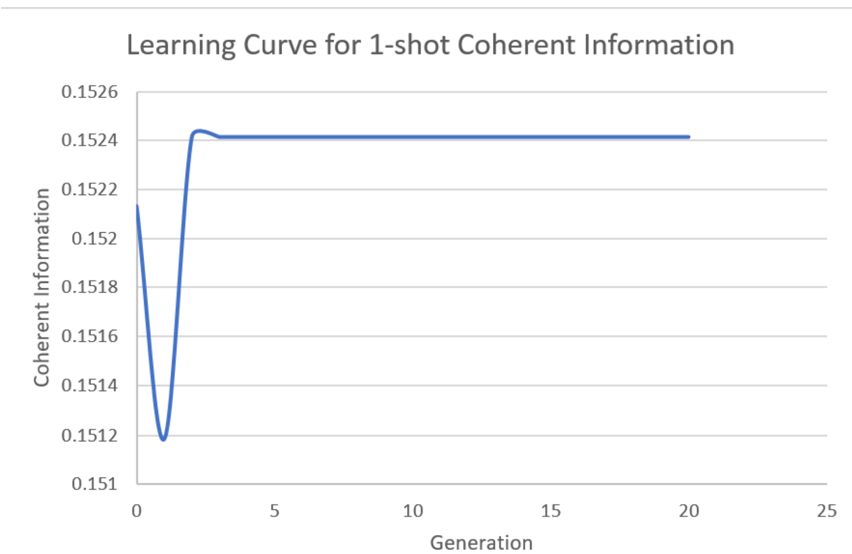


- But when optimization was done for multiple channels, the gradient descent method gave different results for different trials, indicating it might be getting stuck on different local optimas.





- On using the Genetic algorithm as the optimization technique for single channel use of Qubit pauli channels, repetition codes were found as the quantum codes with the highest coherent information.



Pauli Qubit Channel

$|\psi\rangle =$

$0.35357231 + 0.353572i|00\rangle + 0.35357231 + 0.353572i|01\rangle +$
$0.35357231 + 0.353572i|10\rangle + 0.35357231 + 0.353572i|11\rangle$

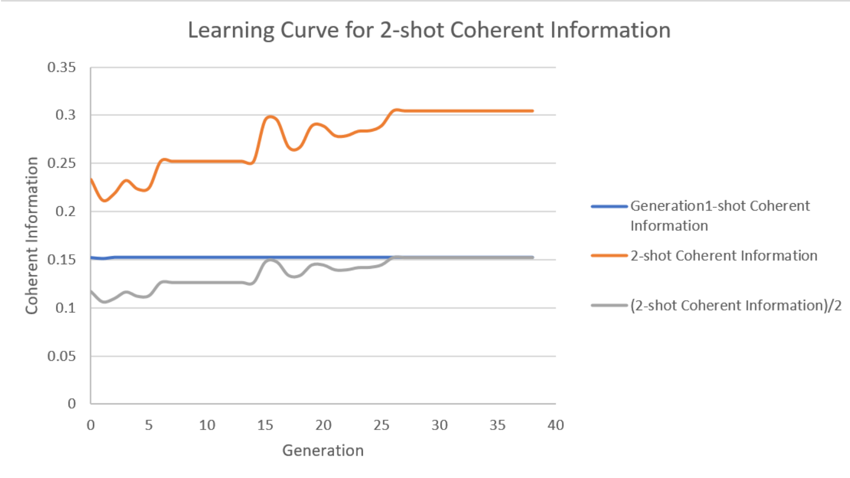
After applying Hadamard to the state introduced for purification we get the repetition code. $\alpha_0|00\rangle + \alpha_1|11\rangle$

- The reliability of using variational state ansatz and genetic algorithm as a possible optimization technique to tackle the issue of local optima in the cases of multiple channels uses of Pauli qubit channel and Dephrasure channels was confirmed.

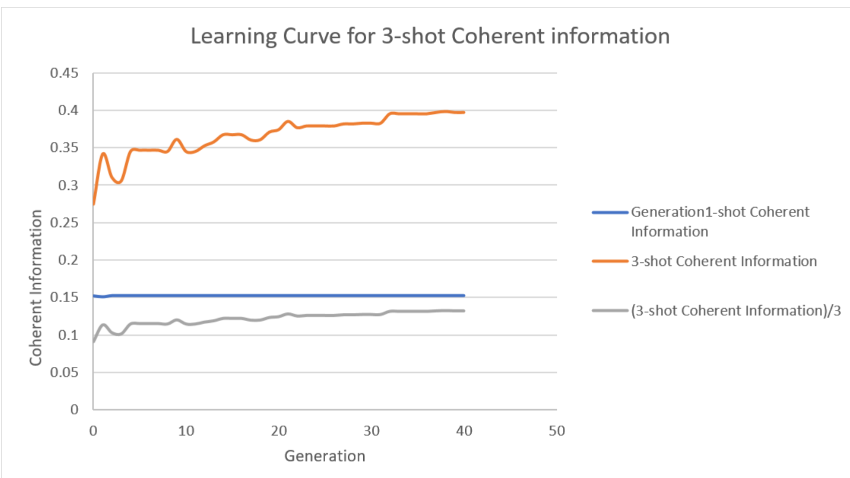
- The simple genetic algorithm used proved efficient in showing quantum channels that exhibited superadditivity only in Qubit pauli channels. In other channels such as Dephrasure channel, Depolarising channel, and Generalized amplitude damping channels it was inefficient in finding superadditive nature.

The superadditive nature was exhibited for the 2-channel use of 0.225688,0.00801196,0.0263041 parameters of Qubit Pauli channel.

(Coherent Information maximized overall quantum states for 2 channel use)$\times\frac{1}{2}$
- (Coherent Information maximized overall quantum states for 1 channel use)
$= 0.004639833933552717$



Pauli Qubit Channel



Pauli Qubit Channel

4.2 Future Direction

- The computation for checking superadditivity for one parameter case was computationally demanding for normal computers, so plotting the same for a parameter space was not practical and hence parallelizing the code and running the code on a system with the suitable number of cores and threads will be attempted.

- More Qubit channels will be included and better optimization using advanced genetic algorithms will be attempted.
- Results seem to show the genetic algorithm still getting stuck on local optima, so improving the algorithm will help in getting a better parameter space that shows superadditivity
- Other evolutionary optimizations scheme will also be explored in order to achieve a benchmark case.
- Attempts to improve the parameter search technique will also be tried, to make the code less computationally demanding

Bibliography

1. C. E. Shannon, "A mathematical theory of communication," The Bell System Technical Journal 27, 379–423 (1948).

2. T. M. Cover and J. A. Thomas, Elements of Information Theory (Wiley, 2005).

3. M. A. Nielsen and I. L. Chuang, Quantum Computation and Quantum Information (Cambridge: Cambridge University Press, 2000).

4. C. H. Bennett and S. J. Wiesner, "Communication via one- and two-particle operators on einstein-podolsky-rosen states," Phys. Rev. Lett. 69, 2881–2884 (1992).

5. C. H. Bennett, G. Brassard, C. Crépeau, R. Jozsa, A. Peres, and W. K. Wootters, "Teleporting an unknown quantum state via dual classical and einstein-podolsky-rosen channels," Phys. Rev. Lett. 70, 1895–1899 (1993).

6. C. H. Bennett and G. Brassard, "Quantum cryptography: Public key distribution and coin tossing," Theoretical Computer Science 560, 7–11 (2014).

7. S. S. Bhattacharya, A. G. Maity, T. Guha, G. Chiribella, and M. Banik, "Random-receiver quantum communication," PRX Quantum 2, 020350 (2021).

8. G. Chiribella, M. Banik, S. S. Bhattacharya, T. Guha, M. Alimuddin, A. Roy, S. Saha, S. Agrawal, and G. Kar, "Indefinite causal order enables perfect quantum communication with zero capacity channels," New Journal of Physics 23, 033039 (2021).

9. T. Guha, M. Alimuddin, S. Rout, A. Mukherjee, S. S. Bhattacharya, and M. Banik, "Quantum Advantage for Shared Randomness Generation," Quantum 5, 569 (2021).

10. S. Saha, S. S. Bhattacharya, T. Guha, S. Halder, and M. Banik, "Advantage of quantum theory over nonclassical models of communication," Annalen der Physik 532, 2000334 (2020)

11. Vaisakh M, R. K. Patra, M. Janpandit, S. Sen, M. Banik, and A. Chaturvedi, "Mutually unbiased balanced functions and generalized random access codes," Phys. Rev. A 104, 012420 (2021).

12. S. Agrawal, R. Tarafder, G. Smith, A. Roy, and M. Banik, "Better transmission with lower capacity: lossy compression over quantum channels," (2021)

13. M. M. Wilde, "From classical to quantum shannon theory," arXiv:1106.1445 (2011).

14. B. Schumacher and M. D. Westmoreland, "Sending classical information via noisy quantum channels," Phys. Rev. A 56, 131–138 (1997).

15. A.S. Holevo, "The capacity of the quantum channel with general signal states," IEEE Transactions on Information Theory 44, 269–273 (1998).

16. I. Devetak, "The private classical capacity and quantum capacity of a quantum channel," IEEE Transactions on Information Theory 51, 44–55 (2005).

17. S. Lloyd, "Capacity of the noisy quantum channel," Phys. Rev. A 55, 1613–1622 (1997).

18. P. W. Shor, The quantum channel Capacity and coherent information (Lecture Notes, MSRI Workshop on Quantum Computation, 2002).

19. M. B. Hastings, "Superadditivity of communication capacity using entangled inputs," Nature Physics 5, 255–257 (2009).

20. K. Li, A. Winter, X. Zou, and G. Guo, "Private capacity of quantum channels is not additive," Phys. Rev. Lett. 103, 120501 (2009).

21. G. Smith and J. A. Smolin, "Extensive nonadditivity of privacy," Phys. Rev. Lett. 103, 120503 (2009).

22. G. Smith and J. Yard, "Quantum communication with zero-capacity channels," Science 321, 1812–1815 (2008).

23. J. Oppenheim, "For quantum information, two wrongs can make a right," Science 321, 1783–1784 (2008).

24. G. Smith, J. A. Smolin, and J. Yard, "Quantum communication with gaussian channels of zero quantum capacity," Nature Photonics 5, 624–627 (2011).

25. T. Cubitt, D. Elkouss, W. Matthews, M. Ozols, D.Pérez-García, and S. Strelchuk, "Unbounded number of channel uses may be required to detect quantum capacity," Nature Communications 6 (2015), 10.1038/ncomms7739.

26. J. Bausch and F. Leditzky, "Quantum codes from neural networks," New Journal of Physics 22, 023005 (2020).

27. D. P. DiVincenzo, P. W. Shor, and J. A. Smolin, “Quantum-channel capacity of very noisy channels,” Phys. Rev. A 57, 830–839 (1998).

28. G. Smith and J. A. Smolin, “Degenerate quantum codes for pauli channels,” Phys. Rev. Lett. 98, 030501 (2007).

29. J. Fern and K. B. Whaley, “Lower bounds on the nonzero capacity of pauli channels,” Phys. Rev. A 78, 062335 (2008).

30. F. Leditzky, D. Leung, and G. Smith, “Dephrasure channel and superadditivity of coherent information,” Phys. Rev. Lett. 121, 160501 (2018).

31. S. Russell and P. Norvig, “Artificial intelligence: a modern approach,” (2002).

32. M. Schuld, I. Sinayskiy, and F. Petruccione, “An introduction to quantum machine learning,” Contemporary Physics 56, 172–185 (2015).

33. J. Biamonte, P. Wittek, N. Pancotti, P. Rebentrost, N. Wiebe, and S. Lloyd, “Quantum machine learning,” Nature 549, 195–202 (2017).

34. G. Carleo, I. Cirac, K. Cranmer, L. Daudet, M. Schuld, N. Tishby, L. Vogt-Maranto, and L. Zdeborová, “Machine learning and the physical sciences,” Rev. Mod. Phys. 91, 045002 (2019).

35. R. Iten, T. Metger, H. Wilming, Lídia del Rio, and R. Renner, “Discovering physical concepts with neural networks,” Phys. Rev. Lett. 124, 010508 (2020).

36. K. Hornik, “Approximation capabilities of multilayer feedforward networks,” Neural Networks 4, 251–257 (1991).

37. M. H. Hassoun et al., Fundamentals of artificial neural networks (MIT press, 1995).

38. Z. Lu, H. Pu, F. Wang, Z. Hu, and L. Wang, "The expressive power of neural networks: A view from the width," in Proceedings of the 31st International Conference on Neural Information Processing Systems (2017) pp. 6232–6240.

39. Y. LeCun, B. Boser, J. S. Denker, D. Henderson, R. E. Howard, W. Hubbard, and L. D. Jackel, "Backpropagation applied to handwritten zip code recognition," Neural Computation 1, 541–551 (1989).

40. A. Madry, A. Makelov, L. Schmidt, D. Tsipras, and A. Vladu, "Towards deep learning models resistant to adversarial attacks," (2017).

41. Z. Cai and J. Liu, "Approximating quantum many-body wave functions using artificial neural networks," Phys. Rev. B 97, 035116 (2018).

42. H. Saito, "Solving the bose–hubbard model with machine learning," Journal of the Physical Society of Japan 86, 093001 (2017).

43. M. Abadi et. al., "TensorFlow: Large-scale machine learning on heterogeneous systems," (2015), software available from tensorflow.org.

44. D. P Kingma and J. Ba, "Adam: A method for stochastic optimization," (2014).

45. T. Bäck, D. B. Fogel, and Z. Michalewicz, Evolutionary computation 1: Basic algorithms and operators (CRC press, 2018).

46. F. A. Fortin, F. M. De Rainville, M. A. Gardner, M. Parizeau, and C. Gagné, "Deap: Evolutionary algorithms made easy," Journal of Machine Learning Research 13, 2171–2175 (2012).

47. J. P. Cohoon and W. D. Paris, “Genetic placement,” IEEE Transactions on Computer-Aided Design of Integrated Circuits and Systems 6, 956–964 (1987).

48. M. M. Wilde, Quantum information theory (Cambridge University Press, 2013).

49. J. Bausch and F. Leditzky, “Error thresholds for arbitrary pauli noise,” SIAM Journal on Computing 50, 1410–1460 (2021)

50. C. H. Bennett, D. P. DiVincenzo, J. A. Smolin, and W. K. Wootters, “Mixed-state entanglement and quantum error correction,” Phys. Rev. A 54, 3824–3851 (1996).

51. F. A. Fortin, F. M. De Rainville, M. A. Gardner, M. Parizeau, and C. Gagné, “Deap: Evolutionary algorithms made easy,” Journal of Machine Learning Research 13, 2171–2175 (2012).

52. S. Yu and et. al., “Experimental observation of coherent-information superadditivity in a dephrasure channel,” Phys. Rev. Lett. 125, 060502 (2020).